

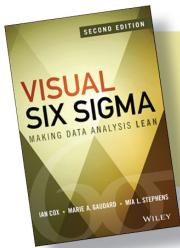
SECOND EDITION

VISUAL SIX SIGMA

MAKING DATA ANALYSIS LEAN

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

WILEY



From *Visual Six Sigma: Making Data Analysis Lean, Second Edition*. Full book available for purchase [here](#).

Contents

Preface to the Second Edition ix

Preface to the First Edition xiii

Acknowledgments xv

About the Authors xvii

PART ONE BACKGROUND 1

- Chapter 1 Introduction 3
- Chapter 2 Six Sigma and Visual Six Sigma 7
- Chapter 3 A First Look at JMP 27
- Chapter 4 Managing Data and Data Quality 67

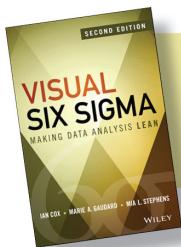
PART TWO CASE STUDIES 101

- Chapter 5 Reducing Hospital Late Charge Incidents 103
- Chapter 6 Transforming Pricing Management in a Chemical Supplier 157
- Chapter 7 Improving the Quality of Anodized Parts 223
- Chapter 8 Informing Pharmaceutical Sales and Marketing 297
- Chapter 9 Improving a Polymer Manufacturing Process 345
- Chapter 10 Classification of Cells 437

PART THREE SUPPLEMENTARY MATERIAL 509

- Chapter 11 Beyond "Point and Click" with JMP 511

Index 539



From *Visual Six Sigma: Making Data Analysis Lean, Second Edition*. Full book available for purchase [here](#).

CHAPTER

9

Improving a Polymer Manufacturing Process

Your employer, a British company called MoldMat Ltd., manufactures granulated white plastic at a plant in Britain and supplies it to a molding plant in Italy, where it is made into white garden chairs and tables. However, the molding process goes through intermittent phases when its product quality drops, leading to yield losses at both the polymer and the molding plants. When a crisis occurs, teams are formed to tackle the problem, but the problem usually disappears for no apparent reason.

After yet another mysterious crisis occurs and resolves itself, you are selected to solve the problem once and for all. Your selection is due in large part to your recent Six Sigma black belt training. Together with a small project team, you identify two characteristics (Ys) that are of paramount importance relative to quality and yield: the polymer's melt flow index (MFI) and its color index (CI).

Together with your team, you reanalyze the most recent crisis team's data. The analysis fails to reveal suspected relationships between the two responses and eight process factors. This leads you to suspect that measurement variation may be clouding results. Consequently, you and your team conduct Measurement System Analysis (MSA) studies on the measured Ys and Xs. The problematic variables turn out to be MFI (one of the two Ys) and filler concentration (one of the Xs).

Once the repeatability and reproducibility issues for these two variables are addressed, your team gathers new data. Your analysis begins with data visualization and exploration. Then it proceeds to modeling relationships. You develop useful models for MFI and CI that include terms that might otherwise have been overlooked, had your team not done extensive data visualization.

You use the **Profiler** to optimize MFI and CI simultaneously. Using sound estimates of the expected variation in the Hot Xs, you simulate the expected distributions for MFI and CI at the optimal settings. The simulations indicate that the parts per million (PPM) rate should be greatly reduced. After running some successful confirmation trials, management implements the changes.

One and a half years later, not a single batch of white polymer has been rejected by the molding plant. The savings from rejected batches alone amount to about £750,000 per annum. Additionally, because there are now no processing restrictions on the molding plant, savings of £2,100,000 per annum are being realized by MoldMat's big customer. This in turn leads to increased sales for MoldMat. These savings come at very little cost, as project-related expenditures were minimal.

Your odyssey takes you and your team through all of the steps of the Visual Six Sigma Data Analysis Process. In particular, you engage in interesting work involving MSAs and modeling using stepwise regression. A list of platforms and options used is given in Exhibit 9.1. The data sets can be found at <http://support.sas.com/visualsixsigma>.

Exhibit 9.1 Platforms and Options Illustrated in this Case Study

Menus	Platforms and Options
Tables	Concatenate
Rows	Exclude/Unexclude Hide/Unhide Colors/Markers Row Selection Clear Row States Data Filter
Cols	New Column Column Info Column Properties Hide/Unhide Exclude/Unexclude
DOE	Full Factorial Design
Analyze	Distribution Histogram Continuous Fit Frequency Distribution Fit Model Standard Least Squares Stepwise All Possible Models Macros—Response Surface Macros—Factorial to Degree Effect Summary Modeling Control Chart Builder Process Capability Analysis Measurement Systems Analysis
Quality and Process	Graph Builder Scatterplot Matrix Surface Plot Profiler Maximize Desirability Sensitivity Indicators Simulator
Graph	
Other	Local Data Filter Column Switcher Non-normal Capability

SETTING THE SCENE

For the past 25 years, MoldMat Ltd. has supplied the plastic that one of its major customers in Italy uses in making white garden furniture. Over the years, mysterious crises occur during which the flowability of the plastic leads to low yields for both MoldMat and their Italian customer. To date, all efforts to find the root cause of these crises have failed. After the most recent crisis, due to your training in Visual Six Sigma, you are asked to lead a new team whose mandate is to find a permanent solution to the problem.

Manufacturing Process

White garden chairs and furniture command a very good price, but they are difficult to make owing to the impact of whitening agents on plastic flowability. Getting the right balance of whiteness and flow is not easy. As the proportion of additives in the mix increases to make the plastic whiter, the flow of the plastic is impaired.

The process for making white plastic begins with the preparation of a filler mixture, or *slurry*. The white filler, which is an inert powder, is sourced from a number of quarries in Africa. It is mixed with unpurified river water in a stirred tank in the filler preparation section of the MoldMat polymer plant (Exhibit 9.2). The filler preparation tank is agitated and held at a target concentration. The tank is topped off each day with filler and water. Small amounts of a viscosity modifier are added to the slurry if the viscosity gets too high.

Clear plastic is made by heating and stirring a monomer in a batch reactor until it polymerizes. To make the white plastic, the filler slurry is added to the monomer in the polymerization reactor at the start of the polymerization process. When the polymerization reaction is complete, the molten polymer is granulated and packed. The MoldMat plant in England makes three batches of white plastic per day, running a 24-hour schedule every day of the week.

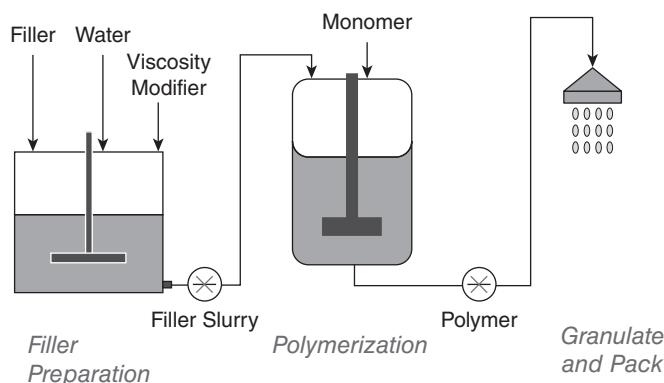


Exhibit 9.2 White Polymer Manufacturing Process

The polymer plant tests every batch of polymer. A sample from each completed batch is taken and tested for:

- Color (whiteness), measured on a colorimeter using a color index
- Melt flow, measured as a melt flow index in an offline laboratory test; this is an indicator of how well the polymer will process in the downstream molding plant
- Filler content of the polymer

A Typical Crisis

Crises have occurred two or three times a year ever since the new product was introduced ten years ago. Here is a typical sequence of events.

The Italian molding plant has several months of normal processing before starting to experience problems with flowability. When this happens, technicians in the molding plant check the processing parameters, and if these look reasonable, they question the quality of the polymer. The MoldMat plant engineers check that the polymer is in specification and verify that there is nothing wrong with the test equipment. This leads the processing plant engineers to suspect that the molding processing parameters have changed.

After a few more days of bad processing, the molding plant engineers ask for some different polymer to run as a trial. This requires a fresh start for molding production. The molding plant must empty the polymer silos to run the trial polymer. The purged material is sold as scrap, which is accounted for as a loss in the MoldMat plant yield.

By this time, the output of the molding plant is well behind schedule, and customers are running out of chairs. The business suffers substantial lost margin and opportunity.

Meanwhile, rapid action teams have been assembled from across Europe. A plethora of helpful theories and their associated solutions are developed, such as:

- The filler supplier is inconsistent and should be replaced.
- Last week's heavy rain has altered the pH of the water supply, which has affected the reaction chemistry.
- The MFI specification is too high, so batches of polymer at the bottom end of the specification range should be the only ones used.
- Abnormal ambient temperatures and humidity are to blame.
- The filler is not evenly distributed through the polymer, and agglomerates are blocking the flow channels in the molds.

Process changes are made, trials are run, and data are gathered. But none of the changes ever conclusively solve the problem.

Then, mysteriously, the problem goes away. The molding process gradually improves, with everyone convinced that their pet theory or solution was the one that made the difference. All is well until the next crisis.

Forming a Team

After one particularly bad crisis, the manufacturing director, Edward Constant, has finally had enough. MoldMat has started to implement Visual Six Sigma, and the black belts from the first wave of training are anxious to start driving improvement. Edward is skeptical about Visual Six Sigma, but he is prepared to give it a go—after all, nothing else has worked.

You are a bright young process engineer who has only recently moved to the polymer plant, and you were one of the first trainees. Edward has met you a few times and is impressed by your openness to new ideas and your approach to problem solving. Given the numerous false starts, Edward figures that your lack of detailed knowledge of MoldMat's operations could actually be an advantage, provided that you work with people who have the right mix of experience.

At your first meeting, Edward indicates that he will act as project sponsor and offers you all the support you need. He tells you: "Everyone has an opinion on the best solution, but I have never been satisfied that anyone has properly done any rigorous analysis, let alone identified the root cause of the problem so that it can be conclusively fixed. This problem has been around for ten years, so a few more months are not going to make that much difference. The best advice I can give you is to take your time and to trust nothing and no one, unless you have personally verified the data and have worked through it in a methodical way. I don't want any more crises. If the process can work most of the time, then it should be able to work all of the time."

Edward knows that a change of polymer can immediately affect the processing performance of the molding plant, even if the polymer batches meet the polymer specifications. So he urges you to focus on the polymer plant first and to talk to a wide range of people in both the polymer and molding plants. But above all, he directs you to collect some data.

Edward helps you form a small project team consisting of you and the following associates:

- Henry, the polymer plant quality manager
- Bill, a polymer chemist from a technical support group
- Roberto, a process engineer from the Italian molding plant
- Tom, a master black belt

Tom's role is to ensure that the Visual Six Sigma methodology and tools are correctly applied. He is a well-seasoned and culturally savvy master black belt

from MoldMat's training partner. Together with Tom, you assemble the team and review your objectives. To ensure that all team members share a common language and approach, Tom schedules and conducts an impromptu training session in Visual Six Sigma and JMP.

FRAMING THE PROBLEM

At this point, the team needs to develop a project charter, using a high-level process map and some baseline data. Also, you decide to develop a *Suppliers, Inputs, Process, Outputs, and Customers* (SIPOC) map and gather customer input. The result of this process is that you and the team decide to focus on two critical process characteristics: melt flow index and color index.

Developing a Project Charter

During your first team meeting, the team draws a high-level process map (Exhibit 9.3). You also decide to review yield data from both the polymer and molding plants to confirm the size and frequency of the problem.

There have been many arguments about white polymer quality. Although there is a polymer specification, the molding plant has long suspected that it does not fully reflect the true requirements of their process. After a long discussion, the team agrees on the following Key Performance Indicator (KPI) definitions for the project:

Daily yield is calculated as the weight of good polymer divided by the weight of total polymer produced.

Good polymer is polymer that can be successfully processed by the molding plant.

Total polymer produced will include product that fails to meet the polymer plant specifications, plus any polymer that, although meeting polymer plant specifications, is subsequently scrapped or rejected in the molding plant.

The team collects some historical data on daily yield and imports it into a data table named **BaselineYieldData.jmp**. The data table contains two

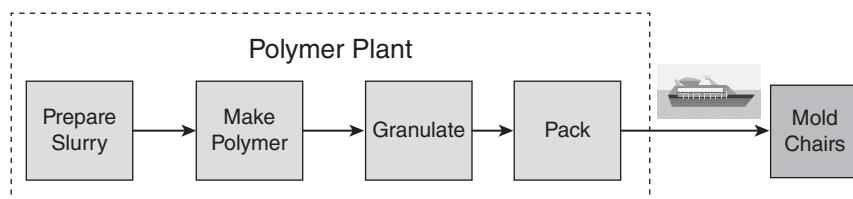


Exhibit 9.3 High-Level Process Map of White Polymer Molding Process

	Date	Yield
1	07/27/2011	100.00%
2	07/28/2011	96.87%
3	07/29/2011	89.42%
4	07/30/2011	95.66%
5	07/31/2011	99.03%
6	08/01/2011	100.00%
7	08/02/2011	93.05%
8	08/03/2011	89.62%
9	08/04/2011	93.25%
10	08/05/2011	96.26%
11	08/06/2011	95.71%
12	08/07/2011	93.61%
13	08/08/2011	92.17%
14	08/09/2011	91.99%
15	08/10/2011	90.07%

Exhibit 9.4 Partial View of **BaselineYieldData.jmp**

columns, *Date* and *Yield*, and has 1,125 rows covering a period of a little more than three years. (The data table also contains two scripts for later use.)

Note that **Yield** is designated as a **Label** variable, as evidenced by the yellow label icon next to *Yield* in the **Columns** panel in Exhibit 9.4. With this property, when the arrow tool hovers over a point in a graph, that point's *Yield* value will appear. You gave *Yield* the **Label** role by right-clicking on it in the **Columns** panel and selecting **Label/Unlabel**.

You proceed to construct an Individuals control chart to see how *Yield* varies over time. The distribution of *Yield* measurements is likely to be skewed, since there is an upper limit of 100 percent, so that control limits calculated using an individual measurement control chart may not be appropriate. Nonetheless, you decide to use the individual measurement chart in an exploratory fashion. You also decide that the Moving Range part of the IR chart is not of interest to you at this point, so you remove it. This allows you and others to focus on the *Yield* values.



To construct an individual measurement control chart for *Yield* (**BaselineYieldData.jmp**, script is **Control Chart Builder**):

1. Select **Analyze > Quality and Process > Control Chart Builder**.
2. Select both *Date* and *Yield* in the **Select Columns** list.
3. Drag and drop these two variables into the center of the template.

4. Right-click on the **Moving Range(Yield)** axis and select **Remove**. (Alternatively, drag the title **Moving Range(Yield)** out of the template area.)
 5. Click **Done** to close the **Control Panel**.
-

The control chart shown in Exhibit 9.5 appears. The chart clearly shows periods of high yields, each followed by a crisis, with a total of nine crises over the period. The average *Yield* over this time period is about 88 percent, but the real issue is the repeated occurrence of causes that are not intrinsic to the system, called *special causes*.

To conveniently reproduce this analysis and to document your work, you can save a script to the data table to reproduce this chart. The script has already been saved for you with the default name **Control Chart Builder**. To save this script yourself, click on the red triangle next to **Control Chart Builder** in the report window and choose **Script > Save Script to Data Table**. By default, your new script is called **Control Chart Builder 2**, but you can click on the name to change it.

To better understand the crisis *Yield* values, use your cursor to hover over various points to see their values. You come to the realization that crisis periods can be loosely defined by collections of batches with *Yield* values below 85 percent.

You become curious about the likely yield of the process had it not been affected by these crisis periods. Just to get a sense of the noncrisis yields, you construct a control chart with crisis batches, defined as batches with yields below 85 percent, excluded. The **Local Data Filter** provides an easy way to filter out data values without affecting other plots or the data table.

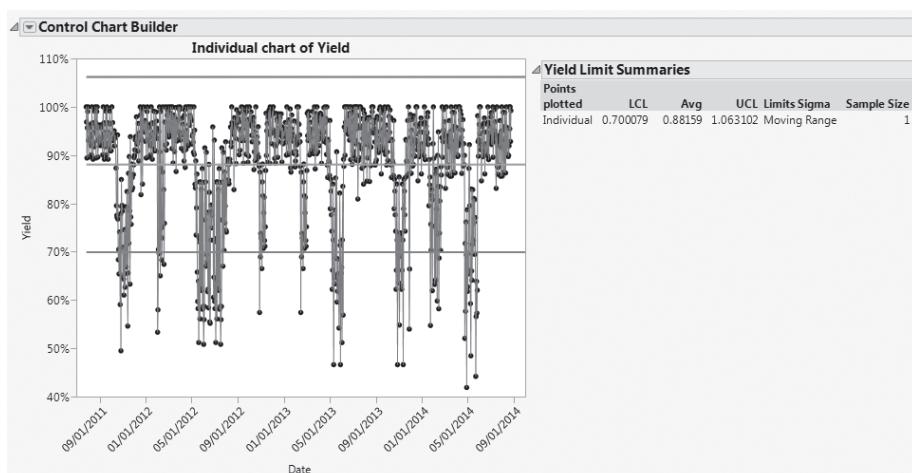


Exhibit 9.5 Individuals Chart of Baseline Yield Data



To filter out the crisis yields (**BaselineYieldData.jmp**, script is **CCB with Local Data Filter**):

1. In the **Control Chart Builder** window, select **Script > Local Data Filter** from the red triangle menu.
2. In the **Local Data Filter** panel, select *Yield* and click **Add**.
3. To include only those rows with yields of at least 85 percent, click on the minimum value for *Yield* above the slider in the dialog box, shown as 41.905, and enter **0.85** into the text box.
4. Click **Enter** (or click elsewhere in the window) to accept the settings.
The control chart updates to show only those rows where *Yield* is at least 85 percent.

The **Local Data Filter** panel and plot are shown in Exhibit 9.6. The Yield Limit Summaries report to the right of the chart, shown beneath the plot in Exhibit 9.6, indicates that the mean *Yield* for the noncrisis periods is about 94 percent. Close your data table without saving changes.

With this information as background, the team reconvenes. You need to agree on the problem statement and project goal, and to define the specific scope and focus of the project.

The team drafts a project charter, shown in Exhibit 9.7. As instructed, you decide to focus on the polymer plant and set a goal of achieving an average yield of 95 percent by the end of the year. It is late August—this gives you four months. You know that if you can eliminate the crises, a 94 percent yield can be achieved. But, knowing that the team will be constructing detailed knowledge of the process, you feel that you can even do a little better.

At this point, you check in with Edward to obtain his support for the project charter and the team's proposed direction. As it turns out, Edward is very impressed with the clarity of your work to date and likes the idea that the project goal was chosen based on sound data. He is quick to approve of the team's charter and direction.

Identifying Customer Requirements

Next, the team explores the following questions:

- What are the true requirements of the molding plant?
- Why are these requirements met at some times but not at others?
- What is changing?

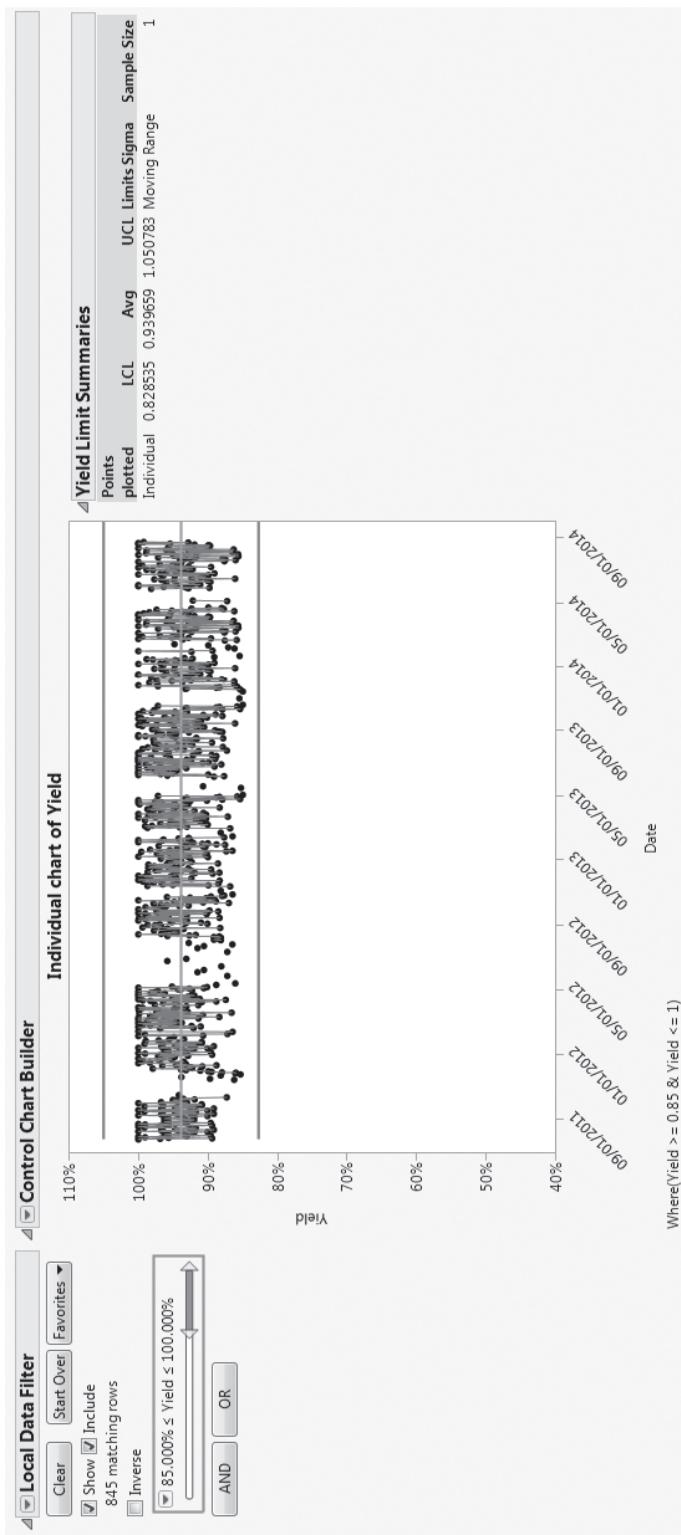


Exhibit 9.6 Local Data Filter Settings to Exclude Crisis Rows

Exhibit 9.7 Project Charter

Project Title	Improve White Polymer Process Yield
Business Case	The manufacture of white polymer results in periodic flowability crises at a large Italian customer's molding plant. The molding plant sells suspect polymer at scrap prices. These crises have been going on for years and, although the crises resolve temporarily, they continue to recur, causing significant disruption and great financial loss for both MoldMat and its customer. Demand for white furniture keeps increasing, and the molding plant in Italy can't afford to be down due to lack of acceptable white polymer. The molding plant has to turn orders away in crisis periods, causing a significant loss in revenue and great dissatisfaction.
Problem/Opportunity Statement	It is estimated that, due to the crisis periods, the polymer plant suffers a yield loss of about £700,000 per year in scrap material. There is the opportunity to recover at least £700,000 annually in what would otherwise be scrap. Also, a significant margin loss is generated by the molding plant, which has to turn orders away in crisis periods. If the problem could be fixed, the accounting department estimates that the company would realize an additional £2,000,000 of revenue annually.
Project Goal Statement and KPI (Key Performance Indicator)	Increase the average yield of white polymer from 88 percent to 95 percent or higher by March 1, 2015 (four months). Daily yield will be plotted using an individual measurement control chart.
Project Scope	The polymer plant's part of the process.
Project Team	Sponsor: Edward Constant Black belt and polymer process engineer: This is you! Team members: Henry Doyle, the polymer plant quality manager Bill Wright, a polymer chemist from a technical support group Roberto Valentino, a process engineer from the Italian molding plant

To this end, during its next meeting, the team produces a SIPOC map to help gain a better understanding of the process steps and to identify where to focus within the polymer plant (Exhibit 9.8).

The team also proceeds to collect *voice of the customer* (VOC) information from the immediate customers of the process, namely, the stakeholders at the molding plant. Through interviews, team members collect information from molding plant technicians and managers. There are many comments reflecting that plant's frustration, such as the following:

- “I don’t want any crises caused by poor polymer.”
- “Your polymer is not consistent.”
- “I don’t believe you when you say you are in spec.”
- “I need to be able to make good white molding all the time.”
- “You are killing my business.”
- “We can’t continue with these scrap levels.”

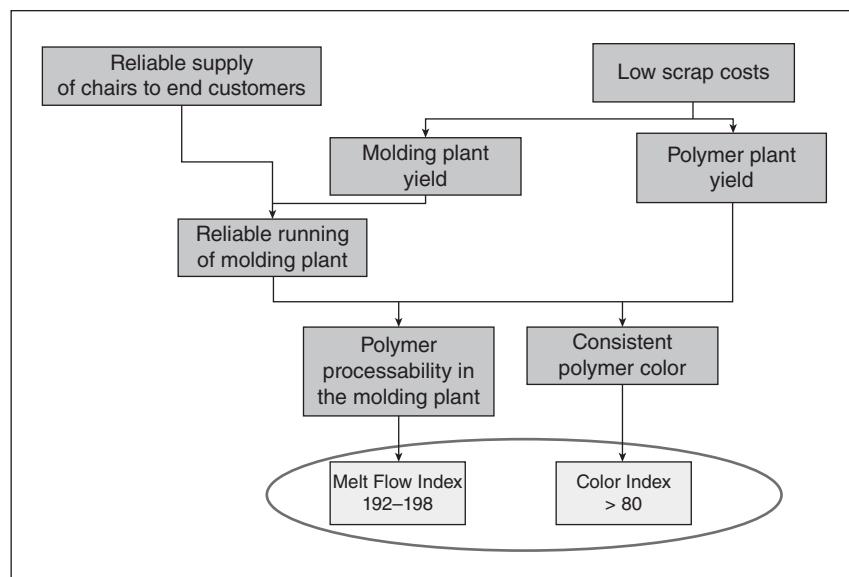
Exhibit 9.8 SIPOC Map for White Polymer Process

Suppliers	Inputs	Process	Outputs	Customers
Umboga A	Filler	Prepare Slurry	Filler Slurry	Polymerization
Kuanga A				
Kuanga B				
North West Water Authority	Water			
Slurry Preparation	Filler Slurry	Make Polymer	Polymer	Granulation
Monomers Inc.	Monomer			
Granulation	Polymer	Granulate	Granules	Packing
Packing	Granules	Pack	Bags	Molding Plant

But the team also collects specific information about the technical requirements of the molding process. You diagram your analysis in the form of a *Critical to Quality Tree*, a portion of which is shown in Exhibit 9.9. This Critical to Quality Tree flows from top to bottom. It first lists the critical customer needs, then their drivers, and finally the measurable requirements needed for improvement.

Two primary characteristics quickly emerge:

- The molding plant has specified that in order for the polymer to process well on their equipment, the polymer's MFI must fall between lower and upper specification limits of 192 and 198 (with a target of 195).
- The polymer's CI must meet the whiteness specification. The maximum possible CI value is 100, but the only requirement is that CI must exceed a lower specification limit of 80.

**Exhibit 9.9** Partial Critical to Quality Tree for Molding Plant VOC

REVIEWING HISTORICAL DATA

Your team decides to review a set of data collected by a team that had been assembled to address one of the recent yield crises. This earlier team was unable to identify the root cause of the crisis. The fact that the team's data do not lead to the identification of Hot Xs leads you to suspect that one or more variables suffers from large measurement error.

Data from Prior Crisis Team

You and your team begin to review the last crisis team's data and analysis. To your surprise and delight, you find that the crisis team had used many Six Sigma tools in investigating possible causes of the problem.

In particular, the team members had developed an *Input/Output* process map (Exhibit 9.10) to help identify the potential Xs that might be driving variation in MFI, CI, and, consequently, Yield. They used the Xs and Ys identified in their process map to determine the data they should collect.

Reanalyzing the Historical Data

You obtain a spreadsheet of the data collected by the crisis team and import this into a JMP table, which you call **CrisisTeamData.jmp**. The data consist of measurements for the Xs and Ys identified in the process map for 127 batches over about a six-week period. A partial view of the data table is shown in Exhibit 9.11.

The columns in the data table are described in Exhibit 9.12. There are three Ys and eight Xs of interest. You note that, even though the table does not have an explicit date or time column, the sequential values of *Batch Number* define the processing order.

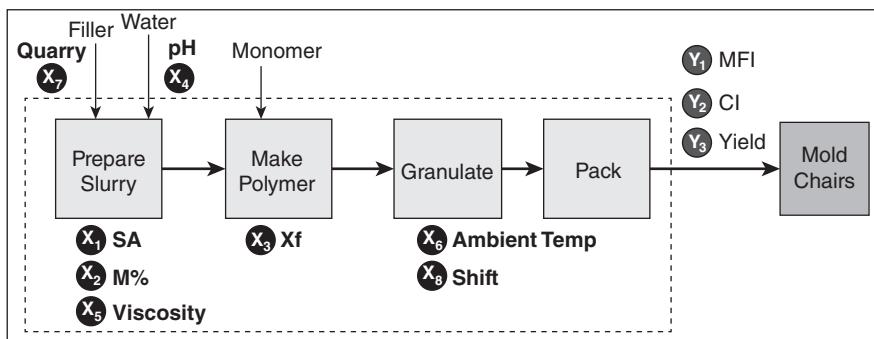


Exhibit 9.10 Input/Output Process Map of White Polymer Process

		Batch Number	MFI	CI	Yield	SA	M%	Xf	pH
	1	3001	213.0	89.2	94.04%	64.6	3.17	14.9	4.70
	2	3002	195.8	86.6	98.50%	65.8	1.98	15.9	5.13
	3	3003	214.3	94.1	98.30%	71.3	1.88	12.8	4.52
	4	3004	189.0	74.1	100.00%	67.9	1.13	22.3	2.75
	5	3005	197.2	98.8	97.72%	64.3	1.13	17.2	5.72
	6	3006	202.5	42.1	97.75%	63.7	1.94	11.7	4.31
	7	3007	196.0	100.0	96.28%	57.0	1.61	17.4	5.41
	8	3008	186.0	98.0	97.11%	64.4	1.54	15.8	3.39
	9	3009	199.4	91.3	94.02%	54.6	2.05	12.5	4.54
	10	3010	190.2	79.5	99.24%	63.8	0.83	14.8	5.13
	11	3011	200.9	93.7	97.87%	56.6	0.81	12.9	4.05
	12	3012	195.9	41.1	100.00%	62.3	0.08	18.1	3.93
	13	3013	205.0	87.0	97.51%	54.9	1.16	15.9	4.32
	14	3014	183.0	91.3	99.70%	65.8	0.19	12.1	4.95
	15	3015	192.8	42.5	99.90%	66.5	0.92	17.7	4.10
	16	3016	206.2	87.0	97.42%	59.9	1.87	18.6	3.41
	17	3017	195.8	98.9	100.00%	78.6	0.29	15.6	3.22
All rows	127	3018	201.5	99.3	97.89%	59.7	1.12	12.6	6.09
Selected	0	3019	199.3	99.6	97.44%	67.2	1.38	14.5	4.68
Excluded	0	3020	201.6	93.7	96.27%	58.6	2.41	18.6	3.35
Hidden	0	3021	204.3	72.8	93.88%	60.1	2.27	9.6	4.10
Labelled	0	--	--	--	--	--	--	--	--

Exhibit 9.11 Partial View of Table Containing Crisis Team Data**Exhibit 9.12** Description of Variables in **CrisisTeamData.jmp**

Variable Type	Name	Description
ID	Batch Number	Identifying number for slurry batch
Ys	MFI	Melt flow index of the polymer
	CI	Color index of the polymer
	Yield	Weight of good polymer as determined by the molding plant, divided by weight of total polymer produced
Xs	SA	Amps for slurry tank stirrer
	M%	Viscosity modifier percent measured in the filler slurry tank
	Xf	Percent of filler in the polymer
	pH	pH of the slurry
	Viscosity	Viscosity of the slurry
	Ambient Temp	Ambient temperature in the slurry tank area
	Quarry	Quarry of origin for filler
	Shift	Shift during which batch was processed

Distribution Plots

The first thing that your team wants to know is how the process behaved, in terms of Ys, over the six-week period reviewed by the previous crisis team. But you first run a **Distribution** analysis for all of the variables (except *Batch Number*). In your training, you learned that this is an important first step in any data analysis.

To obtain a distribution analysis for all variables of interest (**CrisisTeamData.jmp**, script is **Distribution**):

1. Select **Analyze > Distribution**.
2. In the **Select Columns** list, select all columns other than *Batch Number*.
3. Click **Y, Columns**.
4. Click **OK**.

You review the resulting plots. Recall that, in the VOC analysis, you learned that *MFI* should fall between 192 and 198, and *CI* should be 80 or higher. The histogram for *MFI* shows that it varies far beyond the desired limits (Exhibit 9.13). The histogram for *CI* shows a significant percentage of values below 80.

Of particular interest are the *Yield* values that fall below 85 percent. Select these in the box plot to the right of the *Yield* histogram. To do this, click and drag a rectangle that includes these points as shown in Exhibit 9.13.

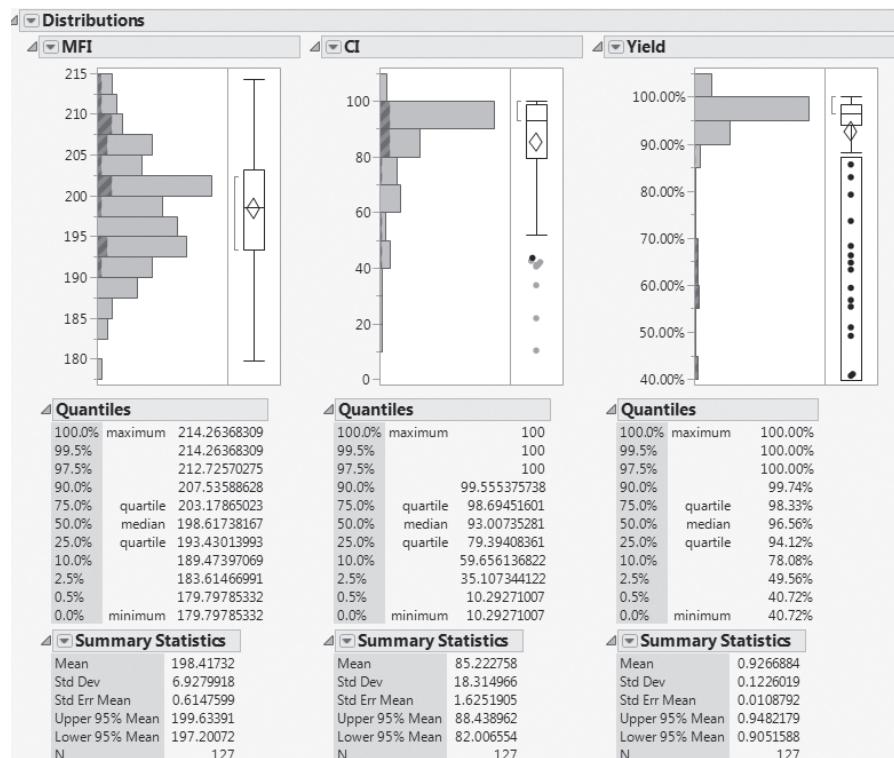


Exhibit 9.13 Partial View of **Distribution** Report with Crisis Yield Values Selected

This action selects the corresponding rows in the data table—check to see that 14 rows have been selected. Consequently, the values corresponding to these 14 low-yielding batches are highlighted in the histograms for all of the variables. For example, you see that the 14 crisis *Yield* batches have very low *SA* values. To remove the selection of the 14 rows, select **Rows > Clear Row States**.

Looking at the histograms, one of the team members points out that the distribution of *CI* is not bell-shaped. You agree and explain that this is not unusual, given that *CI* has a natural upper bound of 100 percent. The team is now aware that they must keep this in mind when using certain statistical techniques, such as individual measurement control charts and capability analysis, both of which assume that measured values are normally distributed.

Control Charts

At this point, you are eager to see how the three Ys behave over time. To see this you construct control charts.



To construct individual measurement control charts for these three responses (**CrisisTeamData.jmp**, script is **Control Chart Builder**):

1. Select **Analyze > Quality and Process > Control Chart Builder**.
2. Drag **Batch Number** to the **Subgroup** area (horizontal axis) in the template.
3. Select **MFI**, **CI** and **Yield** and click **New Y Chart** below the **Select Columns** panel.
Three Individual and Moving Range charts appear.
4. Click **Done** to close the **Control Panel**.

For the most part, *MFI* seems to be stable. On the Individuals chart for *MFI*, there are no points outside the control limits (results not shown). But what is troubling is that *MFI* averages 198.4 over this time period. The team's VOC analysis indicated that 198 is the upper specification limit for *MFI*!

The control chart for *CI* immediately shows the problem with applying a control chart based on normally distributed data to highly skewed data (Exhibit 9.14). The control limits do not reflect the skewed distribution. Nonetheless, there are indications of special causes. The plot also shows some very large and regular dips in *CI*. There are many excursions below the lower specification limit of 80. This leaves the team members puzzled, especially because the dips do not align with the one crisis period that is so evident in the *Yield* control chart (Exhibit 9.15).

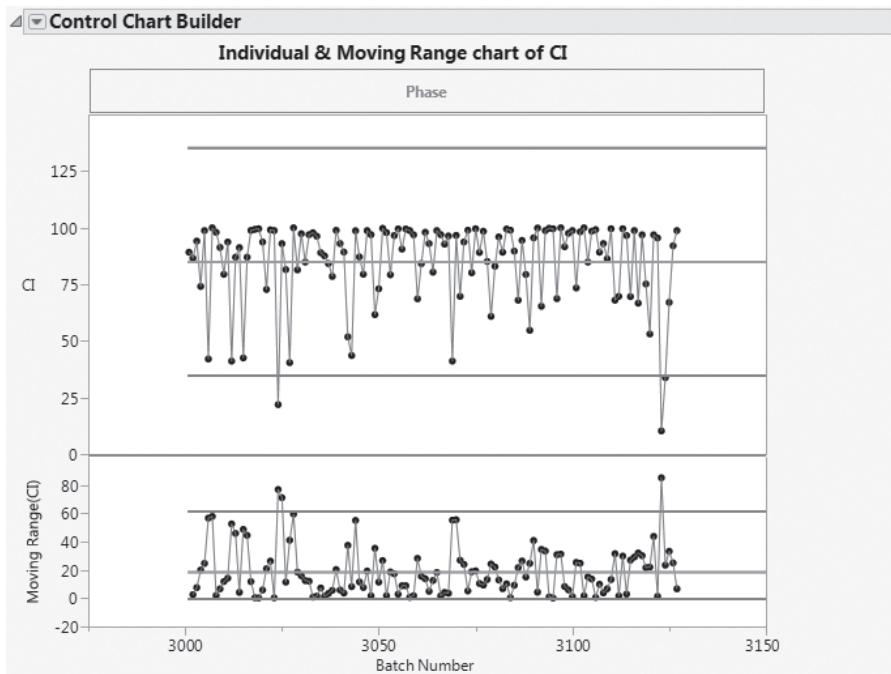


Exhibit 9.14 Individuals Chart for CI, Crisis Team Data

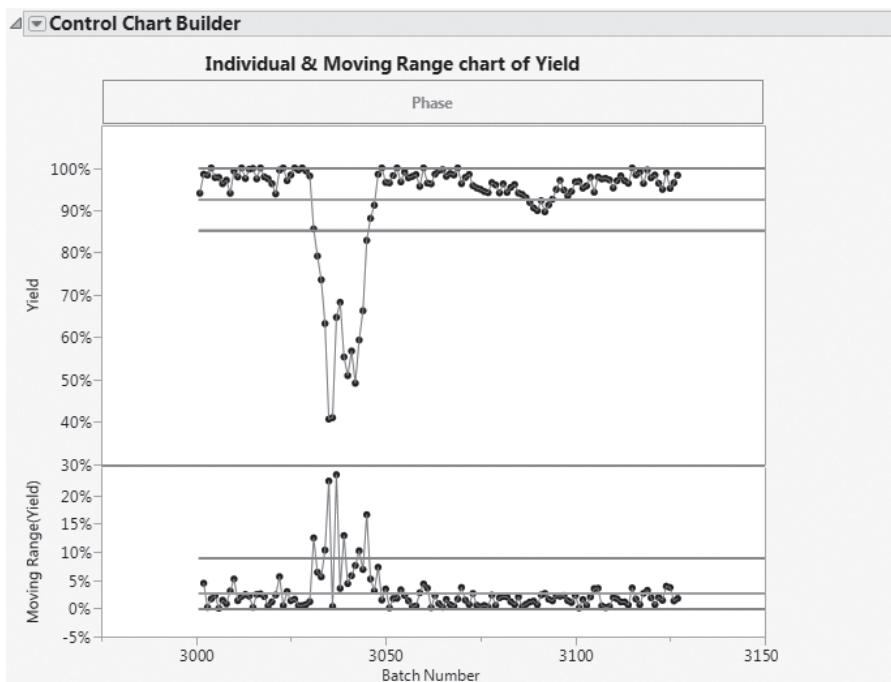


Exhibit 9.15 Individuals Chart for Yield, Crisis Team Data

You briefly consider the idea of running capability analyses for these three responses. However, the Individuals chart clearly shows that *MFI* is not capable, since the process average exceeds the upper specification limit. Each of *Cl* and *Yield* is affected by special causes. The corresponding points could be removed prior to running capability analyses, but it is not all that easy to tell exactly which points result from special causes. The team decides that there is not much to gain by running formal capability analyses on these data.

Now you turn your attention to the crisis team's modeling efforts. Reading through the crisis team's documentation, you see that the team analyzed the data using multiple regression. The crisis team hoped to determine which Xs had a significant effect on the three key responses. The team identified *M%* and *Viscosity* as being significantly related to *MFI*, but it did not find any of the Xs to be related to *Cl*.

This last result seems especially curious. You realize that there are several reasons that such an analysis might lead to no significant factors:

- One or more key Xs are missing.
- One or more higher-order terms involving the specified Xs are missing.
- Measurement variation in the Ys or Xs is too large and is masking the systematic patterns in the Ys caused by process variation in the Xs.

With this as background, you meet with members of the crisis team to discuss the reasoning that led to their final choice of Xs and Ys. They convince you that they did not overlook any critical Xs. After this meeting, you reanalyze their data, introducing higher-order terms. This uncovers some significant relationships, but they don't seem conclusive in terms of the process, especially as they relate to *Cl*. At this point, you have a strong suspicion that measurement variation may be clouding results.

You meet with your team to determine how to proceed. Your team members agree that it is possible that measurement variation in the Ys or Xs could be large relative to the systematic variation caused by the Xs. They fully support your proposal to assess the magnitude of the measurement variation.

MEASUREMENT SYSTEM ANALYSIS (MSA)

As a general principle, measurement systems for key variables should always be evaluated before engaging in data collection efforts. Since you suspect that previous attempts to find root causes of the polymer problems failed because of measurement uncertainty, it is all the more important to thoroughly study the measurement systems for all the variables identified by the process map.

You learn that recent routine MSAs indicate that *Cl*, *SA*, *M%*, *pH*, *Viscosity*, and *Ambient Temp* are being measured with very capable instruments and methods. The same is true of polymer weight, which forms the basis for the *Yield* calculation. However, the measurement systems for *MFI* and *Xf* have not been evaluated in the recent past. Furthermore, given how these measurements are made, you realize that they may be prone to problems.

JMP provides two platforms to perform measurement systems analysis, both found under **Analyze > Quality and Process: Measurement Systems Analysis** and **Variability/Attribute Gauge Chart**. The **Measurement Systems Analysis** platform includes the EMP (Evaluate the Measurement Process) approach.¹ This approach is new to you, and you are anxious to try it out for your team's MSAs.

MSA for MFI

MFI is measured using a melt flow meter during an offline laboratory test. Four instruments are available within the laboratory to perform the test, and there are three different laboratory technicians who do the testing. There is no formal calibration for the instruments. When team members interview the technicians who perform the test, they get the impression that the technicians do not necessarily use a common, standardized procedure.

Properties of a Good Measurement System

You meet with the technicians and their manager to discuss the desired consistency of the measurement process and to enlist their support for an MSA. You discuss the fact that, for characteristics with two-sided specification limits, a guideline that is often used is that the measurement system range, measured as six standard deviations, should take up at most 10 percent of the *tolerance range*, which is the difference between the upper and the lower specification limits. Since the upper and lower specification limits for *MFI* are 198 and 192, respectively, the guideline would thus require the measurement system range not to exceed $10\% \times (198 - 192) = 0.6 \text{ MFI units}^2$.

You also mention that there are guidelines as to how precise the measurement system should be relative to part-to-part, or process, variation: The range of variability of a highly capable measurement system should not exceed 10 percent of the part-to-part (or, in this case, the batch-to-batch) variability.

Designing the MSA

Given these guidelines, you suggest that an MSA for *MFI* might be useful, and the technicians and their manager agree. You learn from the technicians that the test is destructive. *MFI* is reported in units of grams per ten minutes. The protocol calls for the test to run over a half-hour period, with three

measurements taken on each sample at prescribed times, although due to other constraints in the laboratory the technicians may not always be available precisely at these set times. Each of these three measurements is normalized to a ten-minute interval, and the three normalized values are averaged. From preparation to finish, a test usually takes about 45 minutes to run.

Using this information, you design the structure for the MSA. Since the test is destructive, true repeatability of a measurement is not possible. However, you reason, and the technicians agree, that a well-mixed sample from a batch can be divided into smaller samples that can be considered identical. The three technicians who perform the test all want to be included in the study, and they also want to have all four instruments included. You suggest that the MSA should be conducted using samples from three randomly chosen batches of polymer and that the technicians make two repeated measurements for each batch and instrument combination.

This leads to 72 tests: 3 batches \times 3 technicians \times 4 instruments \times 2 measurements. Since each test is destructive, a sample from a given batch of polymer will have to be divided into 24 aliquots for testing. For planning purposes, it is assumed that the MSA design will permit three tests to be run per hour, on average, using three of the four instruments. This leads to a rough estimate of 24 hours for the total MSA. With other work intervening, the technicians conclude that they can finish the MSA comfortably in four or five workdays.

Designing the MSA Experiment

You construct the designed MSA experiment using JMP's design of experiments (DOE) capabilities.



To design the experiment for the MSA (script is **MSA_MFI_Design.jsl**):

1. Select **DOE > Full Factorial Design**.
The resulting window is shown in Exhibit 9.16.
2. Double-click on the response, **Y**, and rename it **MFI**.
3. Click the **Categorical** button under **Factors** to add two three-level categorical factors and one four-level categorical factor.
4. Rename these factors and specify their values as shown in Exhibit 9.17.
5. Click **Continue**.
Note that the default **Run Order** is set to **Randomize**.
6. Insert a value of 1 in the box for **Number of Replicates**, as shown in Exhibit 9.18.
This is because you want two samples to be run at each of these 36 settings.
7. Select **Make Table**.



Exhibit 9.16 DOE Full Factorial Design Dialog

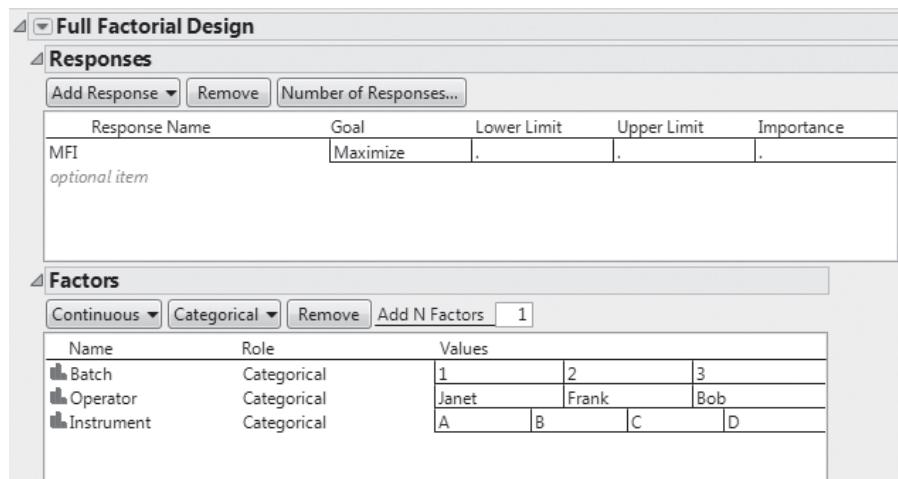


Exhibit 9.17 DOE Full Factorial Design Dialog with Response and Factors Specified

The design table, partially shown in Exhibit 9.19, appears. Most likely, your table will be different because the run order is randomized.

Note that the **Full Factorial Design** dialog remains open. This is useful in case changes need to be made to the design that has been generated. In fact, to obtain the precise run order shown in Exhibit 9.19, go back to the **Full Factorial Design** window. There, select **Set Random Seed** from the red triangle menu, enter **123**, and click **OK**. Then select **Make Table**.

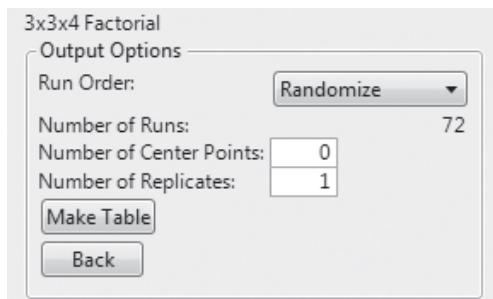


Exhibit 9.18 DOE Full Factorial Design Output Options

	Pattern	Batch	Operator	Instrument	MFI
1	222	2	Frank	B	•
2	234	2	Bob	D	•
3	131	1	Bob	A	•
4	112	1	Janet	B	•
5	314	3	Janet	D	•
6	223	2	Frank	C	•
7	211	2	Janet	A	•
8	232	2	Bob	B	•
9	334	3	Bob	D	•
10	132	1	Bob	B	•
11	311	3	Janet	A	•
12	123	1	Frank	C	•
13	313	3	Janet	C	•
14	111	1	Janet	A	•
15	213	2	Janet	C	•
16	333	3	Bob	C	•
17	314	3	Janet	D	•
18	234	2	Bob	D	•
19	233	2	Bob	C	•

Exhibit 9.19 MSA Design Table

When JMP creates the data table, it automatically includes two data table scripts, **Model** and **DOE Dialog**. The **Model** script specifies a model that you can fit, while the **DOE Dialog** script allows you to recreate the dialog that generated the data table.

Since the purpose of the design is to collect data for a MSA, you decide that the **Model** script is not necessary and so you delete it. You retain the **DOE Dialog** script in case you want to set up another study like this one.

As mentioned earlier, the runs are randomized. You stress the importance of running the trials in this order. The technicians conduct the experiment over the course of the next week and enter their results into the data table.

Conducting the Analysis

Your team regroups to analyze the data. The table that contains the design and results is called **MSA_MFI_Initial.jmp**.



To conduct the MSA analysis (**MSA_MFI_Initial.jmp**, script is **EMP Analysis**):

1. Select **Analyze > Quality and Process > Measurement Systems Analysis**.
Note that **MSA Method** is set to EMP. If you select **Gauge R&R**, you obtain the traditional Variability/Attribute Gauge Chart analysis.
2. Enter **MFI** as **Y, Response**.
3. Enter **Batch** as **Part, Sample ID**.
4. Enter **Operator** and **Instrument** as **X, Grouping**. See Exhibit 9.20.
5. Click **OK**.

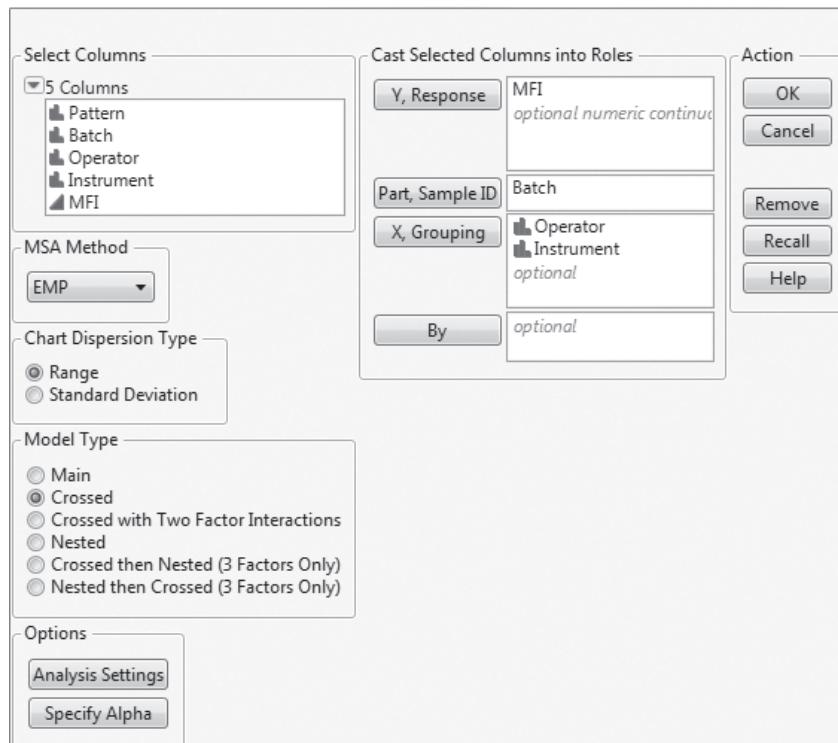


Exhibit 9.20 Launch Window for *MFI* MSA

Interpreting the Results—The Average and Range Charts

The report displays an Average Chart and a Range Chart (Exhibit 9.21). The Average Chart plots the average of the repeated *MFI* measurements and displays control limits calculated from the variation within subgroups. To get a better understanding, select **Show Data** from the Average Chart's red triangle menu to display the individual data points. The control limits are those for an X-bar R chart using the two-measurement subgroups. This means that the limits are for repeatability variation: the variability when the same *Batch* is measured by the same *Operator* using the same *Instrument*.

In a good measurement process, the measurement variability is small relative to the part—in this case, *Batch*—variation. The measurement system allows you to distinguish different parts. For a good measurement process, almost all averages should fall outside the repeatability control limits. But this is not the case here, indicating that the repeatability variation is large compared to the *Batch* variation.

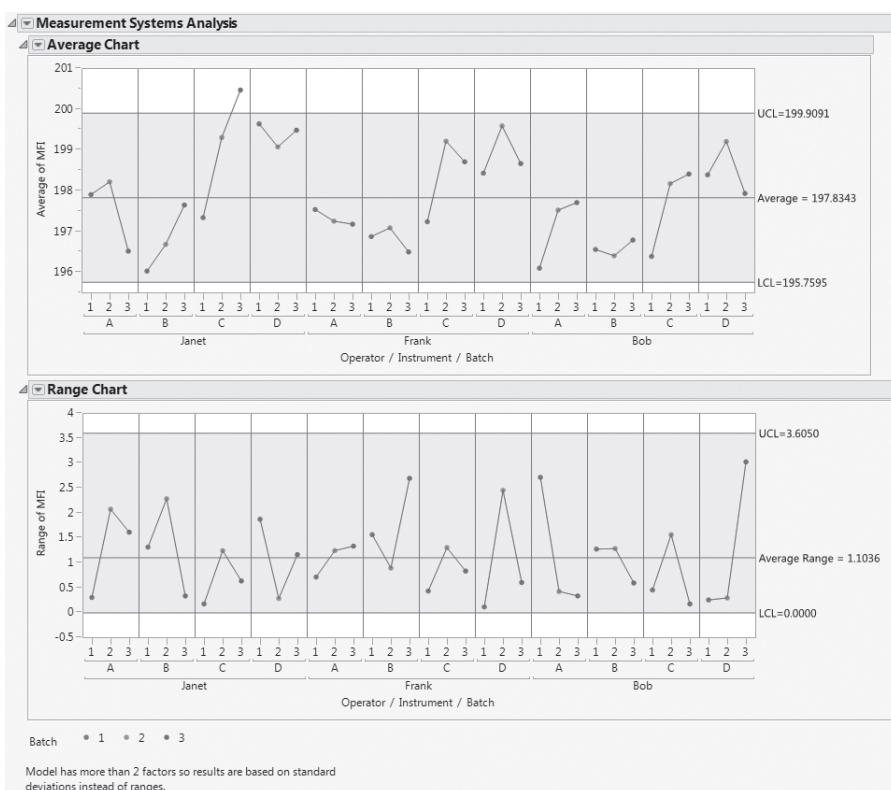


Exhibit 9.21 MSA Average and Range Charts for *MFI* Initial MSA

The Range Chart shows that the repeatability variation is stable. Although this is a positive result, the level at which the process is functioning is not acceptable. The Average Range indicates that the mean difference between two measurements of the same batch by the same operator using the same instrument is 1.1036. The UCL indicates that this could be as large as 3.6050.

Given the fact that the *MFI* measurement system variability should not exceed 0.6 units, this signals a big problem. A sample that falls within the specification limits can easily give a measured *MFI* value that falls outside the specification limits. For example, when Bob used *Instrument D* to measure *Batch 3*, one of his measurements, 196.42 (see row 9 of the data table), indicated that the batch fell within the specification limits (192 to 198). But another measurement, 199.44 (row 62), indicated that it was not within the specification limits. The variability in measurements made on the same batch by the same operator with the same instrument is too large to permit accurate assessment of whether product is acceptable.

Interpreting the Results—Parallelism Plots

For additional graphical insight, select **Parallelism Plots** from the **Measurement Systems Analysis** red triangle menu (Exhibit 9.22). The Parallelism plots show the Mean *MFI* values across *Batch* for both *Operator* and *Instrument*. These plots help you see *Operator* differences across batches, *Instrument* differences across batches, and possible interactions between *Operator* and *Batch* or *Instrument* and *Batch*. Note the following:

- The *Operator* plot suggests that Janet's measurements have higher *MFI* than Bob's. The means of their measurements for the same batches differ by at least 0.5 units.
- It appears that there might be an interaction between *Instrument* and *Batch*. By an *interaction* between *Instrument* and *Batch*, denoted *Instrument*Batch*, we mean that instruments behave differently across batches. Here, Instrument C seems to give higher readings for Batch 2 and Batch 3 than would be expected based on the other instruments.

Interpreting the Results—Variance Components

For a more complete picture of the sources of measurement variation, select **EMP Gauge RR Results** from the **Measurement Systems Analysis** red triangle menu. This report (Exhibit 9.23) shows the standard deviations and variance components associated with the components of the measurement process.

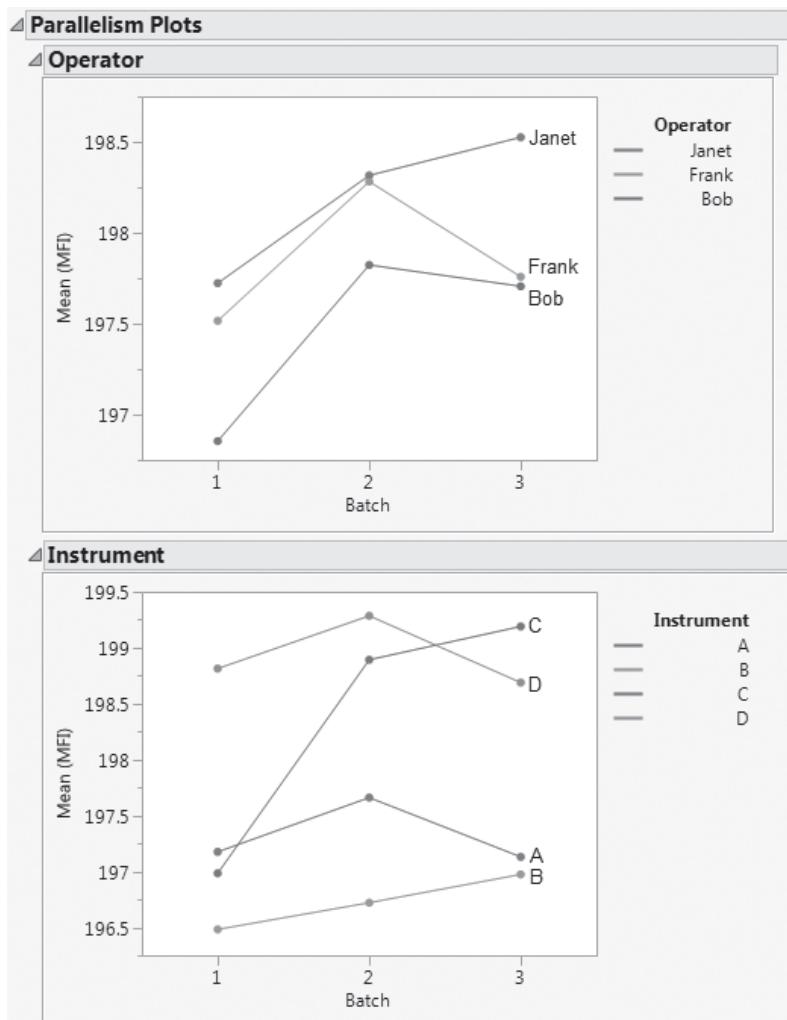


Exhibit 9.22 Parallelism Plots for MFI Initial MSA

EMP Gauge R&R Results

Component	Std Dev	Component	Variance			
			% of Total	20	40	60
Gauge R&R	1.3442331	1.8069627	86.5	86.5	86.5	86.5
Repeatability	0.9124578	0.8325793	39.9	39.9	39.9	39.9
Reproducibility	0.9871086	0.9743834	46.7	46.7	46.7	46.7
Product Variation	0.2953031	0.0872039	4.2	4.2	4.2	4.2
Interaction Variation	0.4402593	0.1938282	9.3	9.3	9.3	9.3
Total Variation	1.4449896	2.0879949	100.0	100.0	100.0	100.0

Exhibit 9.23 EMP Gauge R&R Results for MFI Initial MSA

A *variance component* is an estimate of a variance. The variance components listed in EMP Gauge R&R Results give estimates of the variance in *MFI* values due to:

- *Repeatability*: Repeated measurements of the same part by the same operator with the same instrument.
- *Reproducibility*: Repeated measurements of the same part by different operators using different instruments.
- *Product Variation*: Differences in the parts used in the MSA; here parts are represented by batches.
- *Interaction Variation*: Differences due to the interaction components. Here the interaction components are *Operator*Instrument*, *Operator*Batch*, and *Instrument*Batch*.

The bar graph in the **EMP Gauge R&R Results** panel shows that the Gauge R&R variance component (1.81) is very large compared to the Product Variation (Part-to-Part) variance component (0.09). This suggests that the measurement system has difficulty distinguishing batches. A control chart monitoring this process is mainly monitoring measurement variation. The bar graph also indicates that the repeatability and reproducibility variances, where *Instrument* variation is included in the latter, are large and very close in magnitude, again pointing out that both must be addressed.

The *% of Total* column shows the following:

- Both Reproducibility and Repeatability variation stand out as the large contributors, accounting for 39.9 percent and 46.7 percent of the variation, respectively.
- The combined effect of Reproducibility and Repeatability variation is given by the Gauge R&R component, and is 86.5 percent of the observed variation when measuring the three batches.
- Interaction Variation accounts for 9.3 percent of the total variation.
- The *Batch* variation (Product Variation) accounts for only 4.2 percent of the total variation.

Note the following:

- The variance components for Repeatability and Reproducibility sum to the variance component for Gauge R&R, which is an estimate of the measurement process variance.
- The Total Variation variance component is the sum of the Gauge R&R, Product Variation, and Interaction Variation variance components.
- The individual interaction components are given in the **Variance Components** report, which you can select from the red triangle menu.

- The *Std Dev* values are the square roots of the variance components.
- The *% of Total* is calculated from the *Variance Component* column.
- In an MSA, we are typically not directly interested in part-to-part variation. In fact, we often intentionally choose parts that represent a range of variability.

Recall that the measurement system range, measured as six standard deviations, should take up at most 10 percent of the *tolerance range*. Since the upper and lower specification limits for *MFI* are 198 and 192, respectively, the guideline requires that the measurement system range not exceed $10\% \times (198 - 192) = 0.6 \text{ MFI}$ units.

The *Std Dev* column in the report indicates that the measurement system standard deviation is 1.344. Six times 1.344 is 8.064 MFI units! The measurement variation vastly exceeds the desired precision relative to the specification limits.

Also, the range of variability of a highly capable measurement system should not exceed 10 percent of the Product Variation. Here the *Batch* standard deviation is 0.295 and the Gauge R&R standard deviation is 1.344.

Interpreting the Results—Intraclass Correlation

Select **EMP Results** from the **Measurement Systems Analysis** red triangle menu. This gives you the report in Exhibit 9.24. The *intraclass correlation (with bias and interactions)* is defined as the variance component of Product Variation divided by the variance component of Total Variation. In other words, it is the part variance divided by the sum of the part variance and measurement error variance. When the measurement error is zero, the intraclass correlation is one. When the measurement error variation is large compared to the part variation, the intraclass correlation is small.

The **EMP Results** report indicates that the **Intraclass Correlation** is 0.0418, assuming that you include the reproducibility (*bias*) and interaction variation inherent in the measurement process. In other words, actual *Batch* variation (0.087) is only about 4 percent of the variation that is measured (2.088).

The **EMP Results** report goes on to classify this measurement system as a **Fourth Class** monitor.³ This means that even a large shift in the true *Batch* measurements is unlikely to be detected quickly by a process control chart. For example, the probability of detecting a large three-sigma shift in the true *Batch* values in the next ten subgroups, using the single test of a point beyond the control limits, is very low (between 0.03 and 0.40, see the **Monitor Classification Legend** at the bottom of Exhibit 9.24).

EMP Results	
EMP Test	Results Description
Test-Retest Error	0.9125 Within Error
Degrees of Freedom	36 Amount of information used to estimate within error
Probable Error	0.6154 Median error for a single measurement
Intraclass Correlation (no bias)	0.0948 Proportion of variation attributed to part variation without including bias factors
Intraclass Correlation (with bias)	0.0464 Proportion of variation attributed to part variation with bias factors
Intraclass Correlation (with bias and interactions)	0.0418 Proportion of variation attributed to part variation with bias factors and interactions
Bias Impact	0.0488 Amount by which the bias factors reduce the intraclass correlation
Bias and Interaction Impact	0.053 Amount by which the bias factors and interactions reduce the intraclass correlation

System	Classification
Current (with bias)	Fourth Class
Current (with bias and interactions)	Fourth Class
Potential (no bias)	Fourth Class

Monitor Classification Legend				
Classification	Intraclass Correlation	Attenuation of Process Signal	Probability of Warning, Test 1 Only*	Probability of Warning, Tests 1-4*
First Class	0.80 - 1.00	Less than 11%	0.99 - 1.00	1.00
Second Class	0.50 - 0.80	11% - 29%	0.88 - 0.99	1.00
Third Class	0.20 - 0.50	29% - 55%	0.40 - 0.88	0.92 - 1.00
Fourth Class	0.00 - 0.20	More than 55%	0.03 - 0.40	0.08 - 0.92

* Probability of warning for a 3 standard error shift within 10 subgroups using Wheeler's tests, which correspond to Nelson's tests 1, 2, 5, and 6.

Exhibit 9.24 EMP Results Report for MFI Initial MSA

MSA for X_f

Next, your team enlists technicians in conducting an MSA for X_f . The test that measures X_f , called the *ash test*, measures the concentration of filler in the final polymer as a percent of total weight. The X_f value reported for a given batch is the mean of two or three ash measurements per batch. Adjustments are made to the next batch based on the results of a test on the previous batch. The ash measurement involves taking a sample of polymer, weighing it, heating it in an oven to remove all the organic material, then weighing the remaining inorganic content, which is virtually all filler. The ratio of the filler weight to the initial weight is the reported result and reflects the concentration of filler.

This is a relatively time-consuming test that takes about one hour per sample, so it is imperative that you take this into consideration when designing the study. Also, like the test for *MFI*, the ash test is a destructive test, meaning that repeated measurements will not be true replicates. However, as in the MSA for *MFI*, a sample will be collected from the batch and divided into aliquots. These aliquots will be considered similar enough to form a basis for repeatability estimates.

Designing the MSA

A single dedicated oven is used for the test. Three instruments are available within the laboratory to perform the weight measurements, and there are six different technicians who do the testing. An MSA involving all six technicians would be too time-consuming, so three of the technicians are randomly chosen

to participate. You design a study using samples from three randomly chosen batches of polymer. These will be measured on each of the three instruments by each of the three technicians. Again, two repetitions of each measurement will be taken.

Your design results in a total of 54 tests, with each technician performing 18 tests. At an hour per test, this will take each technician about 18 hours in total; however, the technicians can complete other work while the samples are in the oven. The laboratory manager agrees that this is an acceptable amount of time. The constraint on the duration of the study is the oven, which can be used for only one sample at a time. So, a two-week period is designated for the study, with the intent that *Xf* samples will be worked in between other ash tests performed as part of the laboratory's regular work. To come as close to true repeated measurements as possible, a single sample is taken from each of the three batches of polymer, and the resulting sample is divided into the 18 required aliquots.

You construct the design in a fashion similar to the way you constructed the design for *MFI*. The *Xf* study is completely randomized. Run the script **DOE Dialog** in the data table **MSA_Xf_Initial.jmp** to review how you designed the study.

Conducting the Analysis

The study is conducted and the results are recorded in the data table **MSA_Xf_Initial.jmp**.



Conduct your analysis as you did for *MFI* (**MSA_Xf_Initial.jmp**, script is **EMP Analysis**):

1. Select **Analyze > Quality and Process > Measurement Systems Analysis**.
2. Enter *Xf* as **Y, Response**.
3. Enter *Batch* as **Part, Sample ID**.
4. Enter *Operator* and *Instrument* as **X, Grouping**.
5. Click **OK**.

The Average Chart, shown in Exhibit 9.25, indicates that *Batches* are not distinguished by the measurement process. Individual *Xf* measurements range from about 12 to 18. To see this, select **Show Data** from the Average Chart's red triangle menu. The Range Chart shows that repeated measurements on

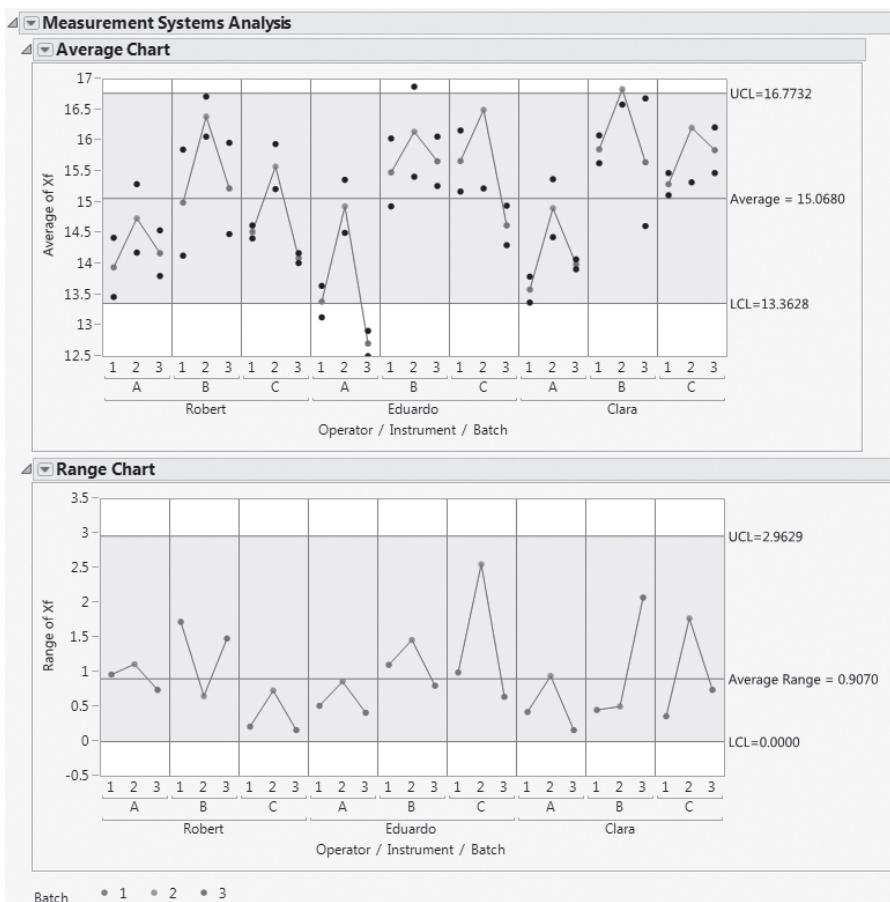


Exhibit 9.25 MSA Average and Range Charts for X_f Initial MSA

the same batch can differ by up to 2.96 units, almost half of the range of the observed data.

The **EMP Gauge R&R Results** are shown in Exhibit 9.26. Repeatability variation is on a par with Product Variation, and Reproducibility variation is more than double the Product Variation.

Your team members observe that, given the variability in readings, it is impossible to differentiate the batches. For example, a team member points out that measurements made by one of the technicians, Eduardo, using Instrument B do not distinguish among the three batches. Repeatability variation seems large as well. Consider Eduardo's two measurements of Batch 2 using Instrument C—click on the point in the Average Chart and view the selected points in the data table. (See Exhibit 9.27.) The two measurements differ by about 2.5 units. Moreover, it appears that measurements made with Instrument A are systematically lower than those made with the other two instruments.

Variance							
Component	Std Dev	Component	% of Total	20	40	60	80
Gauge R&R	1.1278216	1.2719816	73.7				
Repeatability	0.6816677	0.4646708	26.9				
Reproducibility	0.8985048	0.8073108	46.8				
Product Variation	0.6142203	0.3772665	21.9				
Interaction Variation	0.2754972	0.0758987	4.4				
Total Variation	1.3134485	1.7251469	100.0				

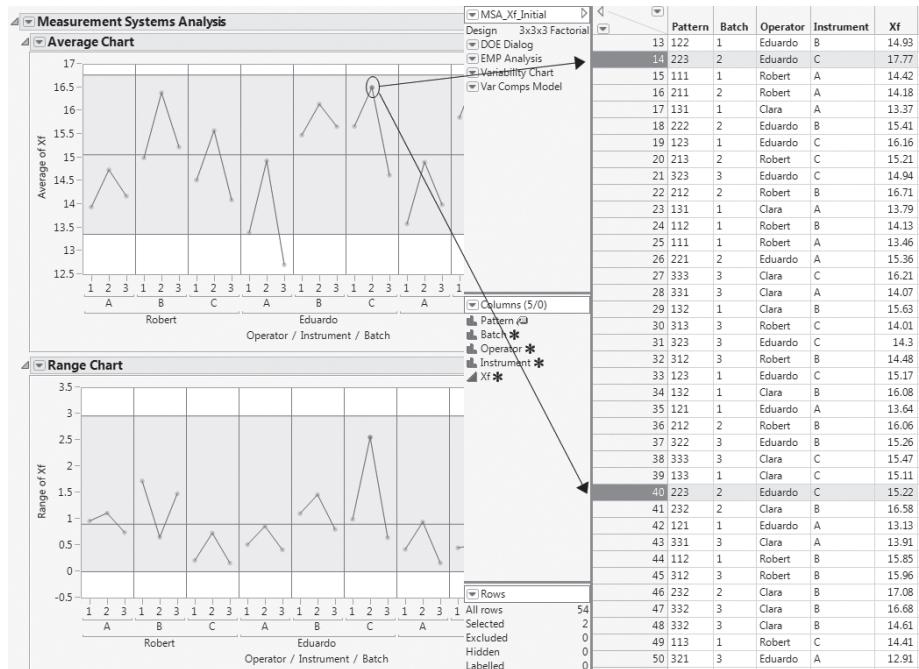
Exhibit 9.26 EMP Gauge R&R Results for X_f Initial MSA

Exhibit 9.27 Eduardo's Batch 2 Measurements with Instrument C

Since no tolerance range for X_f has ever been determined, whether X_f is being measured with enough precision is determined by whether measurements can distinguish different batches, in this case the three batches that were used for the study. (In the production setting, two or three measurements are typically taken, which increases precision. But your intent in this MSA is to estimate the precision of a single measurement.)

The **EMP Gauge R&R Results** indicate that the **Gauge R&R** variance component (1.27) is much larger than the **Product Variation** variance component (0.38). As was the case for **MFI**, this is indicative of a measurement system in trouble. The **Reproducibility** and **Repeatability** variance components are both

large. The **Gauge R&R** standard deviation (1.1278) is almost twice the **Product Variation** standard deviation (0.614).

Select **EMP Results** from the **Measurement Systems Analysis** red triangle menu (Exhibit 9.28). The **EMP Results** report indicates that the **Intraclass Correlation (with bias and interactions)** is 0.2187. This means that actual *Batch* variation accounts for only about 22 percent of the variation that is measured.

The **EMP Results** report also indicates that this measurement system is a Third Class monitor. The probability of detecting a three-sigma shift in the true *Batch* values in the next ten subgroups, using the single test of a point beyond the control limits, is only between 0.40 and 0.88.

Note that the variance components given in the EMP Gauge R&R Results report can also be obtained using **Analyze > Quality and Process > Variability / Attribute Gauge Chart** or using **Fit Model**. In the **MSA_Xf_Initial.jmp** data table, the scripts **Variability Chart** and **Var Comps Model** (see the **Random Effects** tab) illustrate how this is done. Although variances are never negative, sometimes their estimates are negative. Note that, using the script **Var Comps Model**, the variance components are estimated using the *bounded* method so that negative components are set to zero.

Setting a New Timeline and Fixing the Measurement Systems

The findings from the MSAs for *MFI* and *Xf* deal a major blow to the project timetable. But the team can't make progress until these measurement issues are resolved. Furthermore, based on the results of the MSAs, you determine that the historical data obtained by the crisis team are largely useless. Your team needs to collect new data. You estimate that, in total, this will cause at least a three-month delay to the project—six weeks to address the measurement issues and then another six weeks to obtain enough new data to analyze.

You explain the situation to your sponsor, Edward. Despite the projected delay, Edward is very supportive. "So, that means that we have been flying blind for the last ten years. But at least it explains why we have never been able to resolve this issue. Go fix the measurement systems and get some good data and let's see what it tells us!"

With this guidance, you and the team enlist the help of the technicians and engineers who know the measurement processes for *MFI* and *Xf*. Together you begin work on improving the measurement systems.

Fixing the *MFI* measurement process requires that the team address the root causes of the high repeatability variation, reproducibility variation due to instrument, and variation due to the instrument and batch interaction. Observation of the measurement process reveals that, after being removed from the slurry tank, samples can sit for various lengths of time waiting to be analyzed.

EMP Results	
EMP Test	Results Description
Test-Retest Error	0.6817 Within Error
Degrees of Freedom	35 Amount of information used to estimate within error
Probable Error	0.4598 Median error for a single measurement
Intraclass Correlation (no bias)	0.4481 Proportion of variation attributed to part variation without including bias factors
Intraclass Correlation (with bias)	0.2288 Proportion of variation attributed to part variation with bias factors
Intraclass Correlation (with bias and interactions)	0.2187 Proportion of variation attributed to part variation with bias factors and interactions
Bias Impact	0.2193 Amount by which the bias factors reduce the intraclass correlation
Bias and Interaction Impact	0.2294 Amount by which the bias factors and interactions reduce the intraclass correlation

System Classification	
System	Classification
Current (with bias)	Third Class
Current (with bias and interactions)	Third Class
Potential (no bias)	Third Class

Monitor Classification Legend				
Classification	Intraclass Correlation	Attenuation of Process Signal	Probability of Warning, Test 1 Only*	Probability of Warning, Tests 1-4*
First Class	0.80 - 1.00	Less than 11%	0.99 - 1.00	1.00
Second Class	0.50 - 0.80	11% - 29%	0.88 - 0.99	1.00
Third Class	0.20 - 0.50	29% - 55%	0.40 - 0.88	0.92 - 1.00
Fourth Class	0.00 - 0.20	More than 55%	0.03 - 0.40	0.08 - 0.92

* Probability of warning for a 3 standard error shift within 10 subgroups using Wheeler's tests, which correspond to Nelson's tests 1, 2, 5, and 6.

Exhibit 9.28 EMP Results Report for Xf Initial MSA

Also, the analysis steps can occur in various sequences and suffer different time delays. It is suspected that this is a major cause of the repeatability variation. After careful study, a standard operating procedure is developed specifying the sequence of operations and timeline to be followed in processing the samples.

The other issue, reproducibility, revolves around the four melt flow meters. Here, a quick examination of the four meters in use shows that two of them are older units, and that the dies are quite worn. This could account for the differences in how the instruments measure overall, as well as for the batch-specific differences, which are quantified in the *Instrument*Batch* interaction. As part of the improvement strategy, these two units are replaced with new units, and the set of four is tested and calibrated to ensure consistent readings.

Analysis of the measurement process for X_f by your team and technicians also reveals two key issues whose root causes need to be addressed. The first of these is oven control. Sometimes technicians have other tasks that make it difficult to wait until the oven has reached its target temperature or to leave the sample in for the prescribed length of time. This explains the repeatability issues that surfaced in the MSA.

A two-part solution is proposed and approved. First, an oven probe, with a portable remote monitor capable of broadcasting alerts, is purchased. In addition, work assignments are reorganized so that a technician is always in the area of the oven when a test is being run.

The second issue relates to the instrument reproducibility problem. Your team learns that the scales that were being used to weigh the filler were archaic analog scales. They are replaced by high-precision digital scales. Finally, to control other sources of variation, standardized operating procedures for testing X_f are developed and implemented with the help of the technicians.

Follow-Up MSAs for *MFI* and *X_f*

After implementing these improvements to both measurement systems, the team designs follow-up measurement analysis system studies.

Follow-Up for MFI

The follow-up MSA for *MFI* has the same structure as the initial study. The results are given in the table **MSA_MFI_Final.jmp**. The Average and Range Charts are shown in Exhibit 9.29 and the **EMP Gauge R&R Results** are given in Exhibit 9.30. The script is **EMP Analysis**.

The team members and technicians look at the Average Chart and are delighted! Measurements clearly distinguish the three parts. All measurements fall beyond the control limits and, for a given batch, they are very close. The Range Chart indicates that an upper control limit on the range of measurements is 0.4083. There is very little repeatability or reproducibility variation evident.

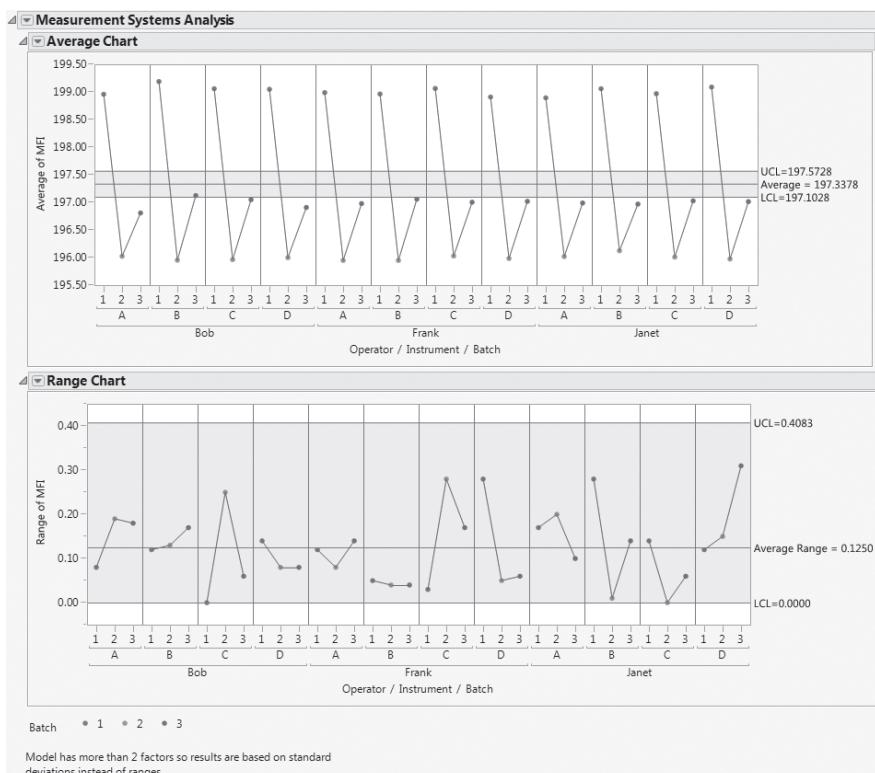


Exhibit 9.29 MSA Average and Range Charts for *MFI* Final MSA

The **Gauge R&R Std Dev**, given in the **EMP Gauge R&R Results** report, is 0.105 (Exhibit 9.30). This means that the measurement system only takes up about $6(0.105) = 0.63$ units, which is almost exactly 10 percent of the tolerance range (recall that the specification limits are 192 and 198). The measurement system is now sufficiently reliable in classifying batches as good or bad relative to the specification limits on *MFI*.

Also, note that the **Intraclass Correlation** (with bias and interactions) is 0.9954 (Exhibit 9.31). This value indicates that there is very little measurement

EMP Gauge R&R Results					
Component	Std Dev	Variance		% of Total	20 40 60 80
		Component	% of Total		
Gauge R&R	0.1050251	0.0110303	0.4645		
Repeatability	0.1011073	0.0102227	0.4305		
Reproducibility	0.0284181	0.0008076	0.034		
Product Variation	1.5373951	2.3635838	99.5		
Interaction Variation	0.0000000	0.0000000	0.0		
Total Variation	1.5409783	2.3746141	100.0		

Exhibit 9.30 EMP Gauge R&R Results for *MFI* Final MSA

EMP Results	
EMP Test	Results Description
Test-Retest Error	0.1011 Within Error
Degrees of Freedom	36 Amount of information used to estimate within error
Probable Error	0.0682 Median error for a single measurement
Intraclass Correlation (no bias)	0.9957 Proportion of variation attributed to part variation without including bias factors
Intraclass Correlation (with bias)	0.9954 Proportion of variation attributed to part variation with bias factors
Intraclass Correlation (with bias and interactions)	0.9954 Proportion of variation attributed to part variation with bias factors and interactions
Bias Impact	0.0003 Amount by which the bias factors reduce the intraclass correlation
Bias and Interaction Impact	0.0003 Amount by which the bias factors and interactions reduce the intraclass correlation

Exhibit 9.31 EMP Results for *MFI* Final MSA

variation relative to batch variation. The **EMP Results** report also indicates that the current system is a First Class monitor. This means that a control chart for *MFI* will signal within ten subgroups with probability of at least 0.99, based on a single point beyond the control limits, if there is shift in the mean that exceeds three standard errors.

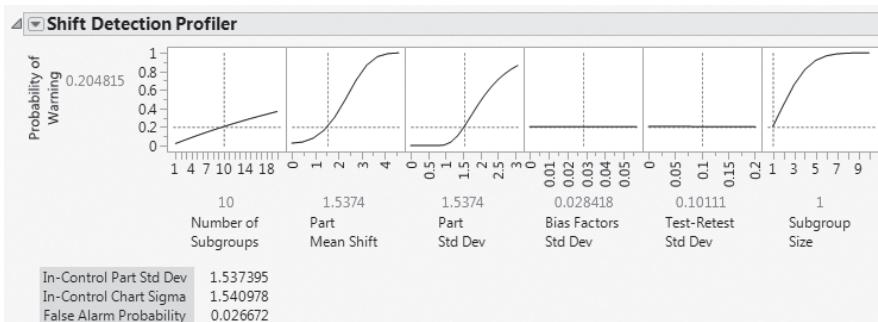
Shift Detection Profiler

You can investigate the sensitivity of a control chart for *MFI* in more detail by selecting the **Shift Detection Profiler** from the topmost red triangle menu in the report. The **Shift Detection Profiler**, which estimates the probability of detecting shifts in the process mean, appears as shown in Exhibit 9.32 (**MSA_MFI_Final.jmp**).

The control limit calculations for the chart include the sources of measurement variation. For the part variation, the control limit calculations use the **In Control Part Std. Dev.** This is initially set to the part standard deviation as estimated from parts used in the MSA. But because parts for MSAs are often not selected at random from process output, you can set the **In Control Part Std Dev** to an appropriate value using a red triangle option.

The **Profiler** shows six cells:

- *Number of Subgroups*: The number of subgroups over which the probability of a warning is computed. This is set to 10 by default.

**Exhibit 9.32 Shift Detection Profiler**—Initial View

- *Part Mean Shift*: The shift in the part mean that you want the control chart to detect. The initial value is set to one standard deviation of the part variation estimated by the MSA analysis. (See **Product Variation** in the **EMP Gauge R&R Results** report, Exhibit 9.30.)
- *Part Std Dev*: The value of the part standard deviation for new points. The initial value is set to one standard deviation of the part variation estimated by the MSA. (See Exhibit 9.30.) You can set the Part Std Dev value to reflect changes in the process.
- *Bias Factors Std Dev*: The standard deviation of factors related to reproducibility, including operator and instrument variability. (See Exhibit 9.30.)
- *Test-Retest Std Dev*: The standard deviation of the test-retest, or repeatability, variation in the model. The initial value is the standard deviation of the Repeatability component estimated by the MSA. (See Exhibit 9.30.)
- *Subgroup Size*: The sample size for each subgroup. This is set to 1 by default.

The initial settings of the **Shift Detection Profiler** are shown in Exhibit 9.32. These settings indicate that, given the current process, the probability of detecting a one standard deviation shift in the mean in the next 10 subgroups, using an individual measurements control chart (Subgroup Size = 1), is about 0.205 (**Probability of Warning** = 0.204815).

What if you were to monitor the process with an Xbar and S chart using a **Subgroup Size** of 5? Slide the vertical bar in the rightmost cell to 5. Then the probability of seeing a warning in the next 10 subgroups is 0.917. (See Exhibit 9.33.)

You can also explore other scenarios, such as the consequences of reducing part or measurement variation. What if you were able to reduce the **In-Control Part Std Dev** to 1.2? Select the option to **Change In-Control Part Std Dev** from the red triangle menu for the **Profiler**. Also, click above **Part Std Dev** in

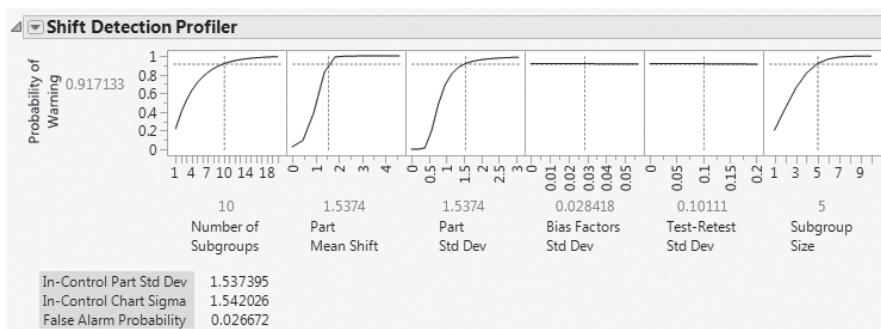


Exhibit 9.33 Shift Detection Profiler, Subgroup Size 5

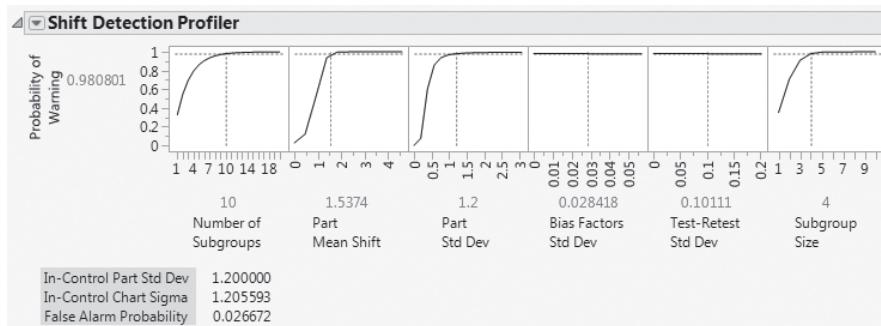


Exhibit 9.34 What-If Scenario for **Shift Detection Profiler**

the third cell and set that standard deviation to 1.2. Change the **Subgroup Size** to 4. (See Exhibit 9.34.) You learn that the probability of detecting the 1.5374 mean shift in the next 10 subgroups, using an X-bar chart based on subgroups of size 4, is about 0.981.

The **Shift Detection Profiler** is a versatile tool, enabling you to interactively explore various scenarios relating to further changes or improvements in how *MFI* is measured, how the process changes, and how the chart itself is constructed. See **Help > Books > Quality and Process Methods** for more details.

Follow-Up for Xf

As for *Xf*, the follow-up MSA is conducted with the three technicians who were not part of the original study. The results are given in **MSA_Xf_Final.jmp**. The script is **EMP Analysis**.

The Average and Range Charts show that the measurement system now clearly distinguishes among the parts. The **EMP Gauge R&R Results** are shown in Exhibit 9.35. The bar graph shows that most of the variation is due to the product.

Once again, the team members and technicians are pleased. Compared to the variability among the three batches, there is very little repeatability or reproducibility variation.

The Gauge R&R Std Dev is 0.034, compared to the Product Variation Std Dev of 0.292. The EMP Results report indicates that the **Intraclass Correlation**

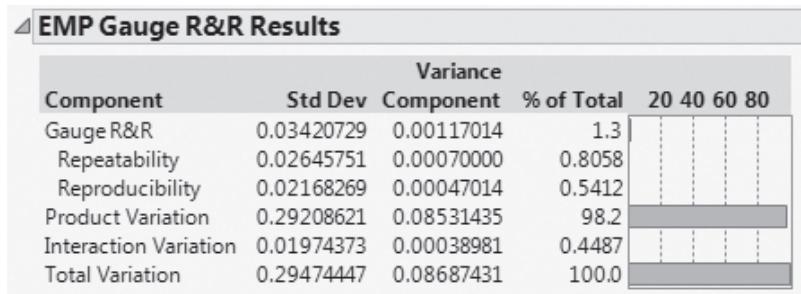


Exhibit 9.35 **EMP Gauge R&R Results** for *Xf* Final MSA

(with bias and interactions) is 0.982, indicating that 98.2 percent of the observed variation is due to part (results not shown).

To ensure that both the *MFI* and *Xf* measurement systems continue to operate at their improved levels, measurement control systems are introduced in the form of periodic checks, monthly calibration, semiannual training, and annual MSAs.

UNCOVERING RELATIONSHIPS

With reliable measurement systems in place, your team now embarks on the task of collecting meaningful process data. The team members collect data on all batches produced during a five-week period. They measure the same variables as were measured by the crisis team, with the assurance that these new measurements have greater precision.

Your analysis plan is to do preliminary data exploration, to plot control charts for *MFI* and *Cl*, to check the capability of these two responses, and then to attempt to uncover relationships between the Xs and these two Ys. You keep your Visual Six Sigma Roadmap, repeated in Exhibit 9.36, clearly in view at all times.

Exhibit 9.36 Visual Six Sigma Roadmap

Visual Six Sigma Roadmap—What We Do	
Uncover Relationships	Dynamically visualize the variables one at a time Dynamically visualize the variables two at a time Dynamically visualize the variables more than two at a time Visually determine the Hot Xs that affect variation in the Ys
Model Relationships	For each Y, identify the Hot Xs to include in the signal function Model Y as a function of the Hot Xs; check the noise function If needed, revise the model If required, return to the Collect Data step and use DOE
Revise Knowledge	Identify the best Hot X settings Visualize the effect on the Ys should these Hot X settings vary Verify improvement using a pilot study or confirmation trials

Visualizing One Variable at a Time

The new data are presented in the data table **VSSTeamData.jmp**.

Distribution Plots

Your first step is to run a **Distribution** analysis for all of the variables except *Batch Number*.

The first five histograms are shown in Exhibit 9.37 (script is **Distribution**). You note the following:

- *MFI* appears to have a mound-shaped distribution, except for some values of 206 and higher.
- *CI* is, as expected, left-skewed.
- *Yield* is also left-skewed, and may exhibit some outliers in the form of low values.

You select those points that reflect low *Yield* values, specifically, those five points that fall below 60 percent, by clicking and drawing a rectangle that includes these points using the arrow tool inside the box plot area. The highlighting in the other histograms in Exhibit 9.38 indicates that these five very low *Yield* rows correspond to four very high *MFI* values, one very low *CI* value, and generally low *SA* values. This is good news, since it is consistent with knowledge that the crisis team obtained. It also suggests that crisis yields are related to Ys, such as *MFI* and *CI*, and perhaps influenced by Xs, such as *SA*. Interestingly, though, four of these rows have *CI* values that exceed the lower specification limit of 80.

A striking aspect of the histograms for *MFI* and *CI* is the relationship of measurements to the specification limits. Recall that *MFI* has lower and upper specification limits of 192 and 198, respectively, and that *CI* has a lower specification limit of 80.

From the **Quantiles** panel for *MFI*, notice that all 110 observations exceed the lower specification limit of 192 and that about 50 percent of these exceed the upper specification of 198. From the **Quantiles** panel for *CI*, notice that about 25 percent of *CI* values fall below the lower specification of 80. To see how often both *MFI* and *CI* meet the specification limits, you use the Data Filter.



To select the points for which the specifications on *MFI* and *CI* are jointly met (**VSSTeamData.jmp**, script is **Data Filter**):

1. Select **Rows > Data Filter**.
 2. Select *MFI* from the **Add Filter Columns** list and click **Add**.
 3. Click on the maximum value of 209.350 above the *MFI* slider and change it to 198.
This selects all rows where *MFI* meets the specification limits.
 4. Click on **AND** in the bottom left corner of the **Data Filter** dialog.
 5. Select *CI* from the **Add Filter Columns** list and click **Add**.
 6. Click on the minimum value of 21.90 above the *CI* slider and change it to 80.
 7. Click **Enter** (or click outside the text box).
-

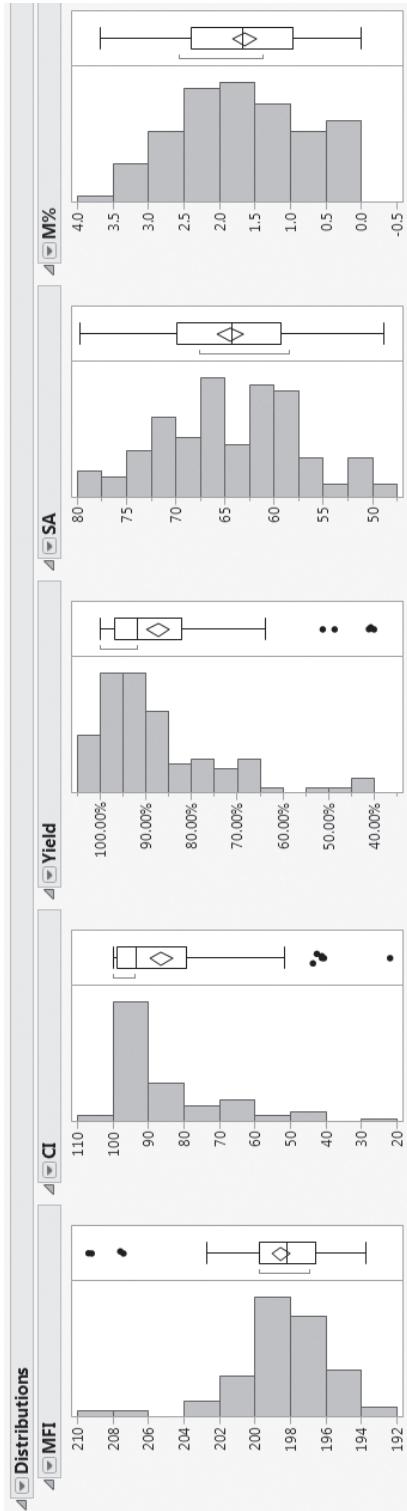


Exhibit 9.37 Five of the Eleven Distribution Reports

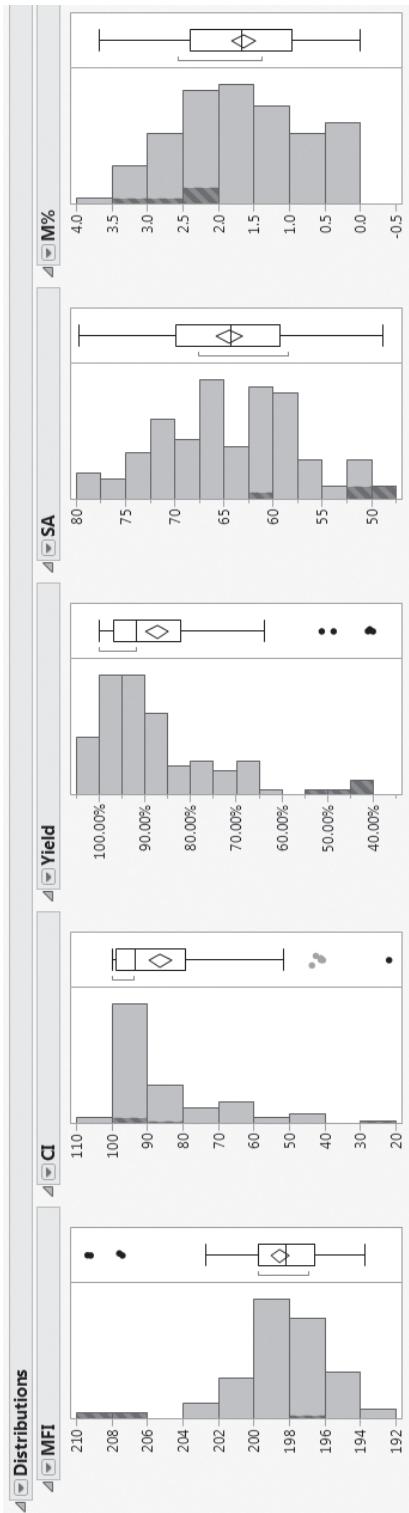


Exhibit 9.38 Distribution Reports with Five Crisis Yield Values Selected

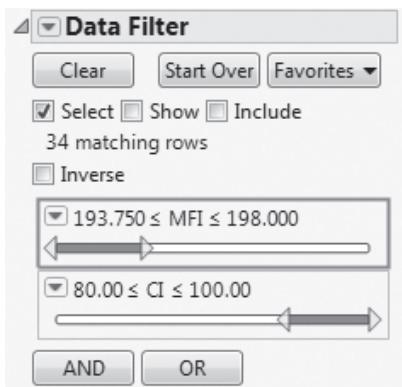


Exhibit 9.39 Data Filter Dialog to Select In-Specification Batches

The completed **Data Filter** dialog is shown in Exhibit 9.39. The **Data Filter** panel indicates that only 34 rows match the conditions you have specified. This means that only 34 out of 110 batches conform to the specifications for both responses.

All rows that meet both specification limits are highlighted in all of the plots obtained using the **Distribution** platform (Exhibit 9.40). As you review the plots, you notice that the batches that meet the joint *MFI* and *CI* specifications tend to result in *Yield* values of 85 percent and higher. They also tend to be in the upper part of the *SA* distribution and the lower part of the *M%* and *Ambient Temp* distributions. This suggests that there might be relationships between the Xs and these two Ys.

Before you close the **Data Filter** dialog, you click the **Clear** button at the top of the dialog to remove the **Select** row states imposed by the **Data Filter**.

Control Charts

With this as background, the team proceeds to see how the three Ys behave over time. As before, you construct individual measurement charts for these three responses using the **Control Chart Builder**. But, this time, you start by constructing an Individuals chart for *MFI* and then use the **Column Switcher** to construct charts for *CI* and *Yield*.

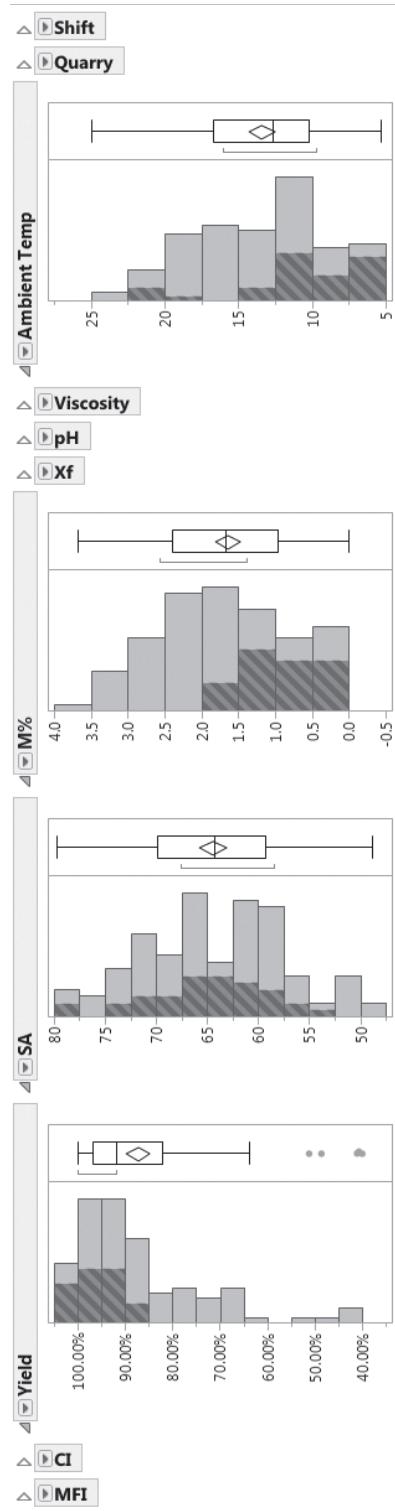


Exhibit 9.40 Points That Meet Both MFI and CI Specifications

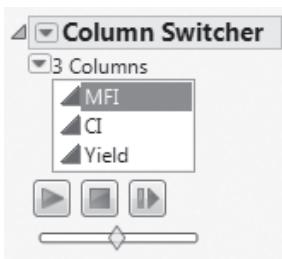


Exhibit 9.41 Column Switcher

To use the **Column Switcher** from the *MFI* individual measurement control chart to plot *CI* and *Yield* (**VSSTeamData.jmp**, script is **Control Charts**):

1. Create the individual measurement control chart for *MFI* as shown earlier.
2. From the red triangle menu next to **Control Chart Builder**, select **Script > Column Switcher**.
3. From the **Choose column to switch** list, select *MFI* and click **OK**.
4. From the **Choose set of columns to switch to** list, select *MFI*, *CI*, and *Yield* and click **OK**.
5. Select the Ys from the **Column Switcher** list to view each control chart in turn. See Exhibit 9.41.

These control charts are quite informative (see Exhibit 9.42). You learn that the average yield is roughly 87 percent, and that there are five batches below the lower control limit. Four of these are consecutive batches. You see that the grouping of four batches with low yields greatly exceed the upper control limit for *MFI*.

From the control chart for *CI* you learn that the fifth yield outlier, corresponding to row 14, falls below the lower control limit. However, you are reminded that the control limits for *CI* are suspect because of the extreme non-normality of its distribution.

You decide that it might be a good idea to assign markers to the five batches that are *Yield* outliers in order to identify them in subsequent analyses.

To mark these five batches in future analyses (**VSSTeamData.jmp**, script is **Markers for Outliers**):

1. Make sure that these five points are selected in the *Yield* control chart.
2. Select **Rows > Markers** and choose square markers.

3. Select **Rows > Colors** and choose a bright red color for the square markers.
 4. Deselect the selected rows by clicking in an empty part of the control chart.
-

Exhibit 9.42 shows the three control charts with the markers you constructed for the five crisis batches.

Visualizing Two Variables at a Time

One of your team members observes that the *MFI* and *Yield* control charts gave advance warning of the crisis in the form of a trend that began perhaps around batch 4,070. Had *MFI* been monitored by a control chart, this trend might have alerted engineers to an impending crisis. You agree but observe that this assumes a strong relationship between *MFI* and *Yield*.

This is as good time as any to see if there really is a strong relationship between *MFI* and *Yield* and *CI* and *Yield*. You also want to explore other bivariate relationships.

Relationships between the Ys

MFI and *Yield*

To explore the relationship between *MFI* and *Yield*, you use **Graph Builder**.

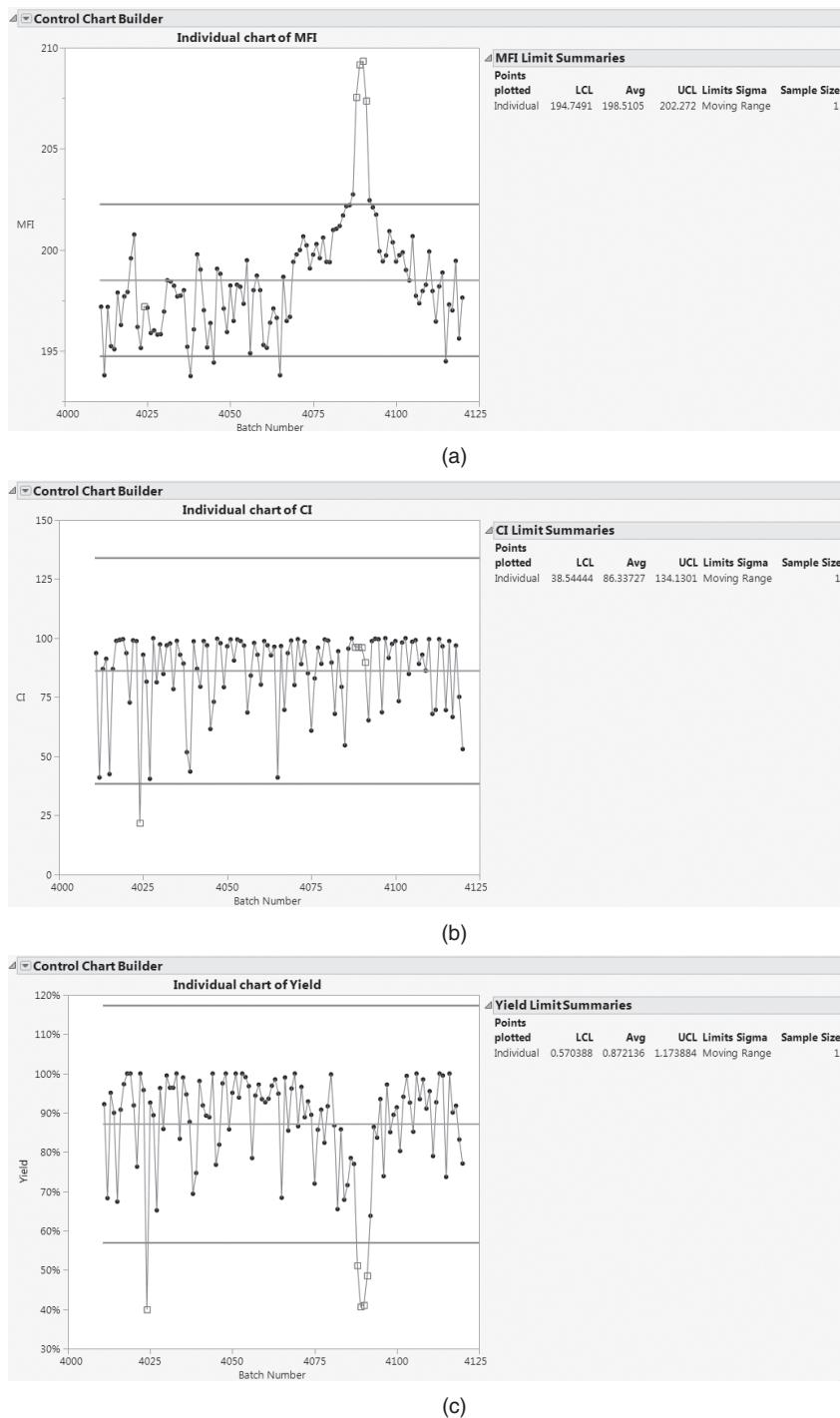


To investigate the relationship between *MFI* and *Yield* (**VSSTeamData.jmp**, script is **Graph Builder—Yield and MFI**):

1. Select **Graph > Graph Builder**.
 2. Select *MFI* from the **Variables** list and drag it to the **X** zone.
 3. Select *Yield* from the **Variables** list and drag it to the **Y** zone.
 4. Click **Done** to close the Control Panel.
-

The points and a default smoother are plotted (Exhibit 9.43). If you wish, you can deselect the smoother by clicking the second icon from the left at the top of the plot before you click **Done** (or, right-click on the graph and select **Smoother > Remove**).

The plot suggests that there is a strong nonlinear relationship between these two Ys. Four of the five outliers clearly suggest that high *MFI* values are associated with crisis level yields. You recognize the outlier with a *Yield* of about 40 as row 14, the point with the low *CI* value identified on the control chart for *CI*. (If you hover over this point, its row number, *MFI*, and *Yield* will appear.)

**Exhibit 9.42** Control Charts with Markers

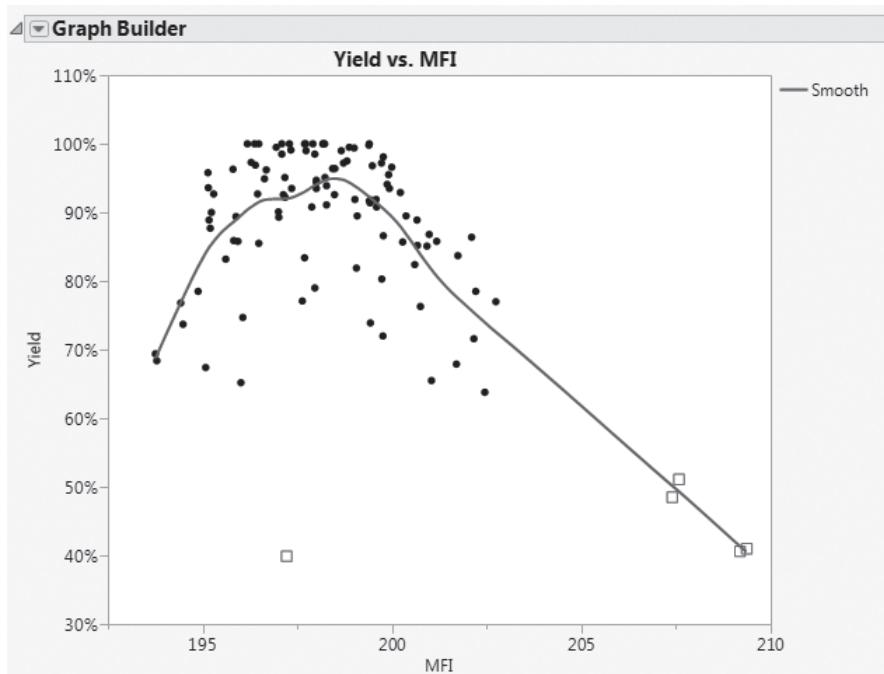


Exhibit 9.43 Graph Builder Plot of Yield by MFI

CI and Yield

You follow the instructions above, replacing *MFI* with *CI*, to construct a plot of *Yield* versus *CI* (see Exhibit 9.44—the script is **Graph Builder—Yield and CI**).

The plot shows a general tendency for *Yield* to increase as *CI* increases. Four of the five outliers have high values of *CI*. The outlier with the low value of *CI* is row 14.

MFI and CI

Now you investigate the relationship between *MFI* and *CI*. You construct the plot shown in Exhibit 9.45, replacing *Yield* with *MFI* (the script is **Graph Builder—MFI and CI**).

MFI and *CI* seem to have a weak positive relationship. There may be process factors that affect both of these Ys jointly. In the next section, we look at relationships between the process factors and these responses.

Relationships between Ys and Xs

As a more efficient way to view bivariate relationships, you decide to create a scatterplot matrix of all responses (Ys) with all factors (Xs), shown in Exhibit 9.46.

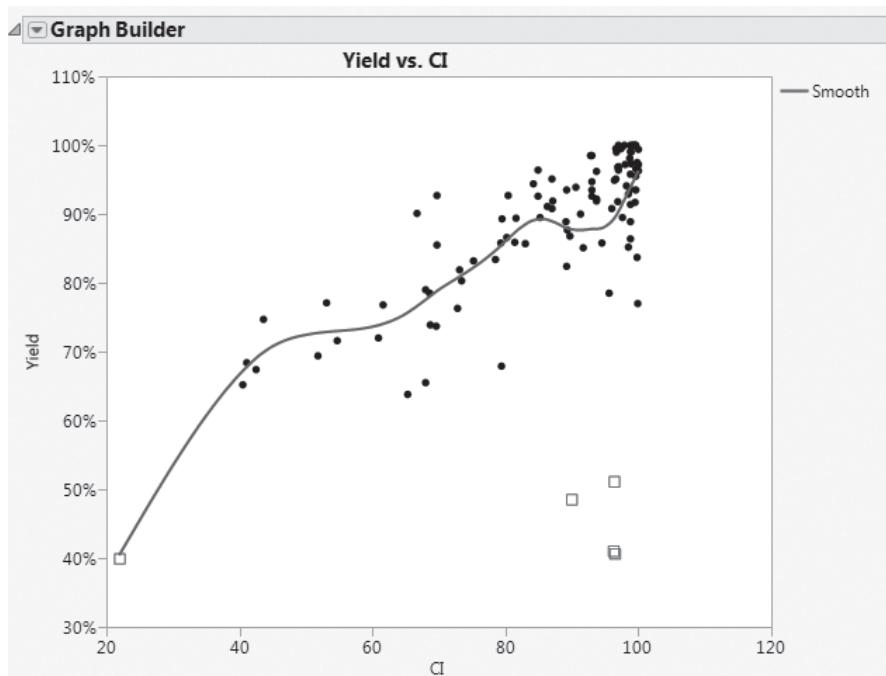


Exhibit 9.44 Graph Builder Plot of Yield by CI

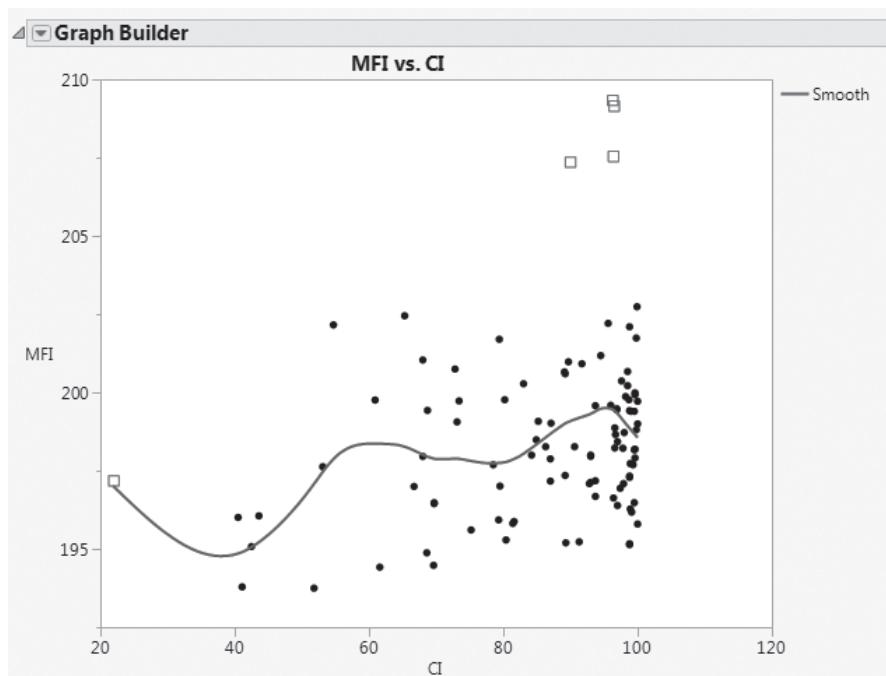


Exhibit 9.45 Graph Builder Plot of Yield by CI

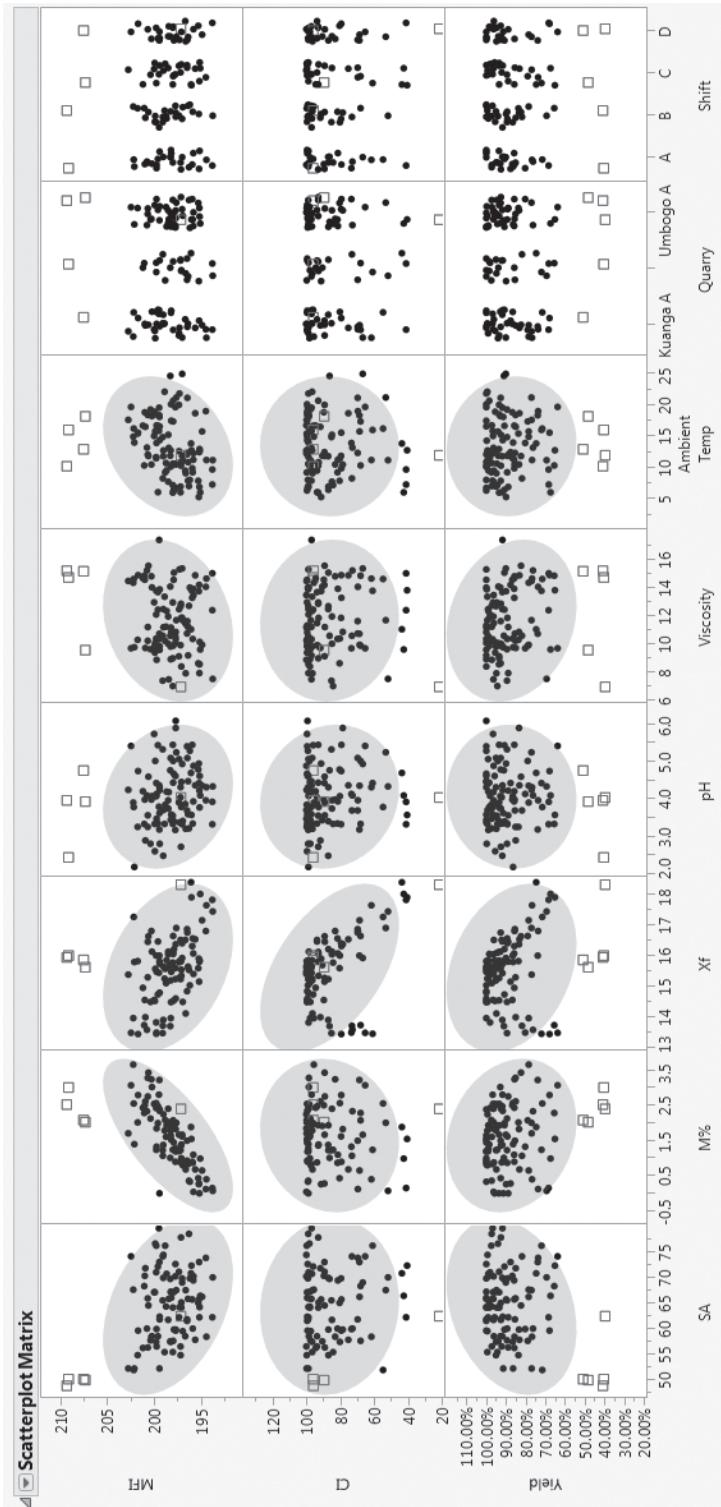


Exhibit 9.46 Scatterplot Matrix of Ys by Xs



To construct the **Scatterplot Matrix** (**VSSTeamData.jmp**, script is **Scatterplot Matrix 1**):

1. Select **Graph > Scatterplot Matrix**.
2. Enter **MFI**, **CI**, and **Yield** as **Y, Columns**.
3. Enter all of the variables from **SA** to **Shift** as **X**.
4. Click **OK**.
5. From the red triangle menu in the report, select **Density Ellipses > Shaded Ellipses**.
6. By dragging the axes, resize them to make each cell include the entire density ellipse.

The matrix shows a scatterplot for each Y and X combination, including scatterplots that involve the two nominal variables, **Quarry** and **Shift**. For the nominal values involved in these scatterplots, the points are jittered randomly within the appropriate level. Density ellipses assume a joint bivariate normal distribution, so they are not shown for the plots involving **Quarry** and **Shift** because these factors are nominal.

The five outliers appear prominently in the **MFI** and **Yield** scatterplots. In fact, they affect the scaling of the plots, making it difficult to see other relationships. To better see the other points, select the five points in one of the **Yield** plots, then right-click in an empty part of the plot and choose **Row Hide and Exclude** from the menu that appears. (Alternatively, run the script **Hide and Exclude Outliers**.) Notice the following:

- Excluding the points causes the ellipses to be automatically recalculated.
- Hiding the points ensures that they are not shown in any of the plots.

Check the **Rows** panel in the data table to verify that the five points are excluded and hidden. Rescale the vertical axes to remove whitespace and accommodate all density ellipses. The updated scatterplot matrix is shown in Exhibit 9.47 (script is **Scatterplot Matrix 2**).

Viewing down the columns from left to right, you see evidence of moderate-to-strong relationships between:

- **MFI** and **SA**
- **MFI** and **M%**
- **MFI** and **CI** (and **Yield**) and **Xf**
- **MFI** and **Ambient Temp**

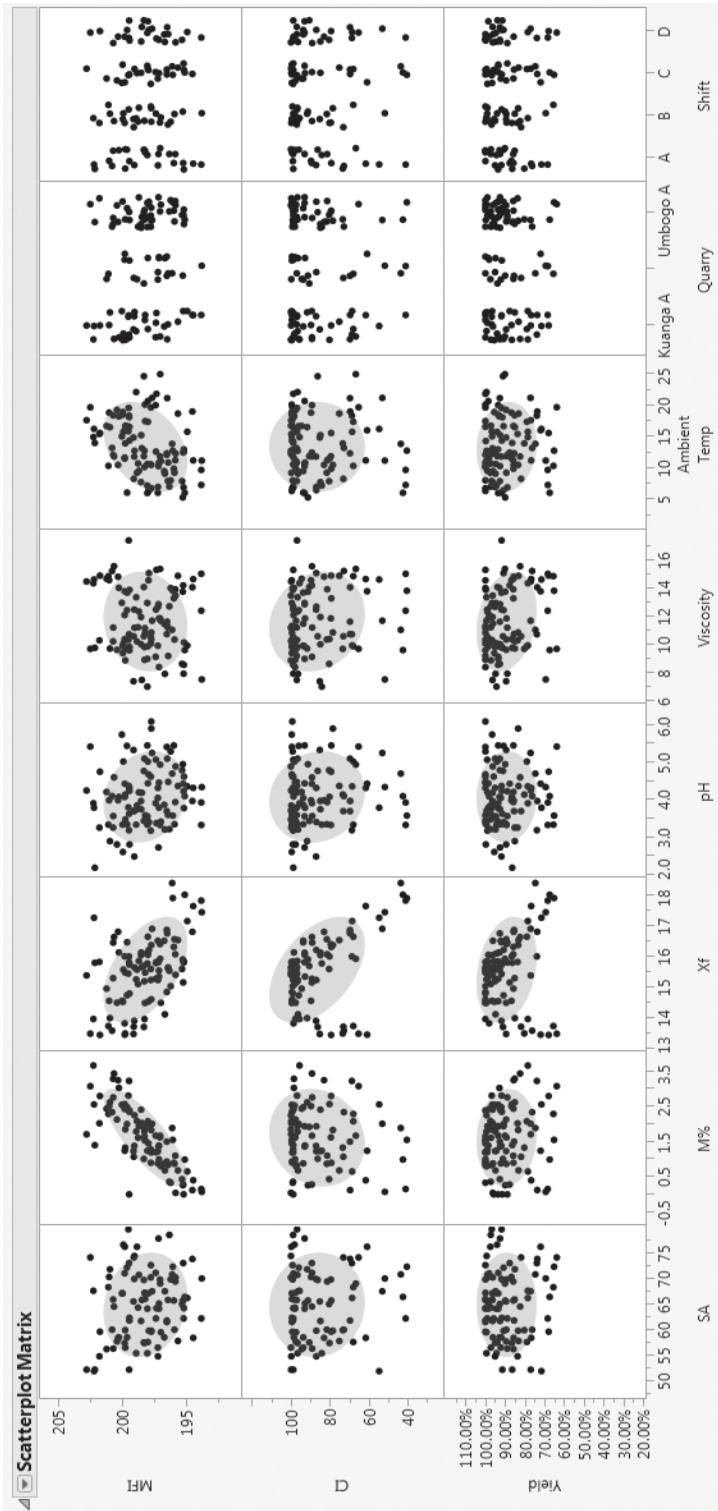


Exhibit 9.47 Scatterplot Matrix of Ys by Xs with Five Outliers Excluded

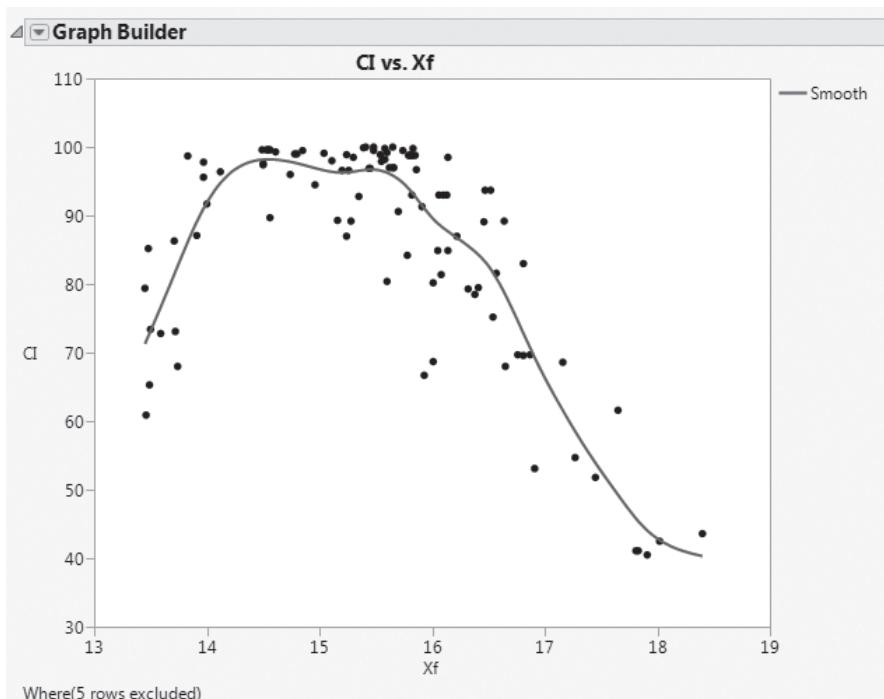


Exhibit 9.48 **Graph Builder** Plot of CI by Xf

The relationship between CI and Xf appears to be highly nonlinear. To see this relationship more clearly, you again use the **Graph Builder** (see Exhibit 9.48—the script is **Graph Builder—CI and Xf**).

Everyone on your team notes the highly nonlinear relationship. An engineer on your team observes that the relationship is not quadratic. She speculates that it might be cubic. You also observe that both low and high values of Xf are associated with CI values that fail to meet the lower specification limit of 80.

The engineer also states that the relationship between MFI and SA might also be nonlinear, based on the underlying science. In a similar fashion, you construct a plot for these two variables using Graph Builder (script is **Graph Builder—MFI and SA**). The plot is shown in Exhibit 9.49.

The relationship does appear to be nonlinear. The bivariate behavior suggests that the underlying relationship might be quadratic or even cubic.

One team member starts speculating about setting a specification range on Xf , maybe requiring Xf to fall between 14.0 and 15.5. You point out that in fact the team must find operating ranges for all of the Xs. You caution that setting these operating ranges one variable at a time is not a good approach. The operating ranges have to simultaneously satisfy specification limits on two

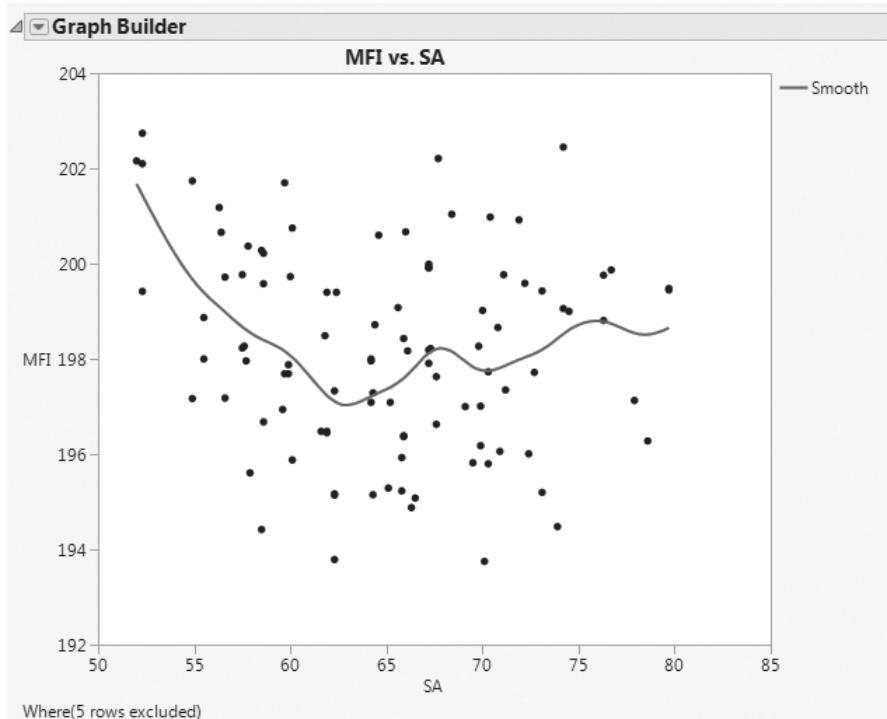


Exhibit 9.49 Graph Builder Plot of MFI by SA

Ys, both *MFI* and *CI*. A statistical model relating the Xs to the Ys would reveal appropriate operating ranges and target settings for the Xs. This brings the team to the *Model Relationships* step of the Visual Six Sigma Data Analysis Process.

MODELING RELATIONSHIPS

The data exploration up to now has revealed several relationships between the two Ys of primary interest, *MFI* and *CI*, and some of the Xs. In particular, *Xf* is related to both *MFI* and *CI* while *SA*, *M%*, and *Ambient Temp* appear to be related to *MFI*. The relationships between *CI* and *Xf* and between *MFI* and *SA* appear to be nonlinear.

At this point, your team embarks on the task of modeling the relationships in a multivariate framework.

Dealing with the Preliminaries

Your first issue involves determining how to deal with the five *MFI* outliers. Your goal is to develop a model that is descriptive of the process operating

under common cause conditions. Your main concern about these five points is whether they are the result of the common cause system that you want to model. If the five points are consequences of a different (special cause) failure mode, their inclusion in the modeling process could result in a less useful model for the common cause process of interest.

The most effective course of action relative to outliers of this kind is to identify and remove the special causes that may have produced them. This is usually best accomplished in real time, when the circumstances that produce an outlier are fresh in peoples' minds.

Looking back at the histograms and control charts for *MFI*, you see that four of these points are very much beyond the range of most of the *MFI* measurements. In fact, they are far beyond the specification limits for *MFI*. The fifth batch is potentially a *CI* outlier. The fact that all five points are associated with extremely low yields also suggests that they are not typical of operating conditions. This evidence seems to suggest that a different set of causes is operative for the five outliers.

Given these findings, you are concerned that the outliers might be detrimental in developing a model for the common cause system. With this as your rationale, you decide to exclude the five crisis observations from the model development process.

However, we suggest the following as an exercise for the reader. Develop models for *MFI* and *CI* that include these five rows. Check to see if the five points are influential, using visual techniques and statistical measures such as Cook's D. Determine how your conclusions would change if you had taken this modeling approach.

Excluding and Hiding the Crisis Rows

The five crisis observations have already been excluded and hidden in connection with your scatterplots for the three responses. You can check the **Rows** panel of the data table to make sure that five points are **Excluded** and **Hidden**. However, if you have cleared row states since then, reselect the outliers as shown in Exhibit 9.50, right-click in the plot, and select **Row Hide and Exclude** (the script is **Hide and Exclude Outliers**).

Saving the Specification Limits as Column Properties

In the interests of expediency for subsequent analyses, you decide to store the specification limits for *MFI* and *CI* in the data table. You do this by entering the information in the **Spec Limits** column property for each column.

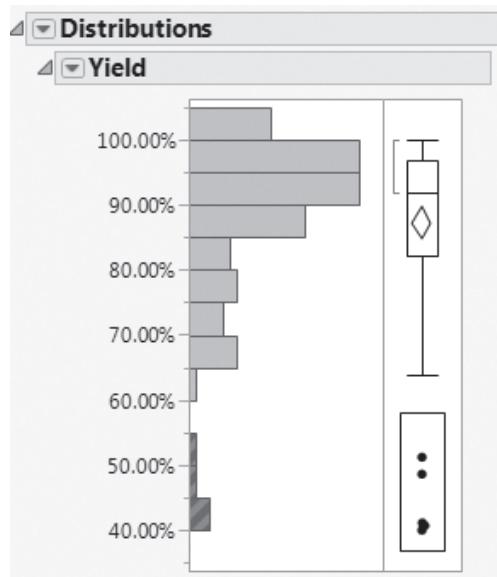


Exhibit 9.50 Selecting the Five Outliers to Hide and Exclude



To save the specification limits for *MFI* as a column property (**VSSTeamData.jmp**, script for both *MFI* and *CI* is **Spec Limits**):

1. Right-click in the *MFI* column header and choose **Column Properties > Spec Limits**.
2. In the **Spec Limits** panel, enter **192** as **Lower Spec Limit**, **195** as **Target**, and **198** as **Upper Spec Limit**.
See Exhibit 9.51. Do *not* click the box next to **Show as graph reference lines**. Seeing the spec limits on control charts can be misleading.
3. Click **OK**.
4. Proceed in a similar fashion for *CI*, recording only a lower specification limit of **80**.

In the **Columns** panel of the data table, you see that asterisks have appeared to the right of *CI* and *MFI*. Click on one of the asterisks to see that the **Spec Limits** property is listed (Exhibit 9.52). If you click on **Spec Limits**, the **Column Info** window opens and displays the **Spec Limit** property panel.

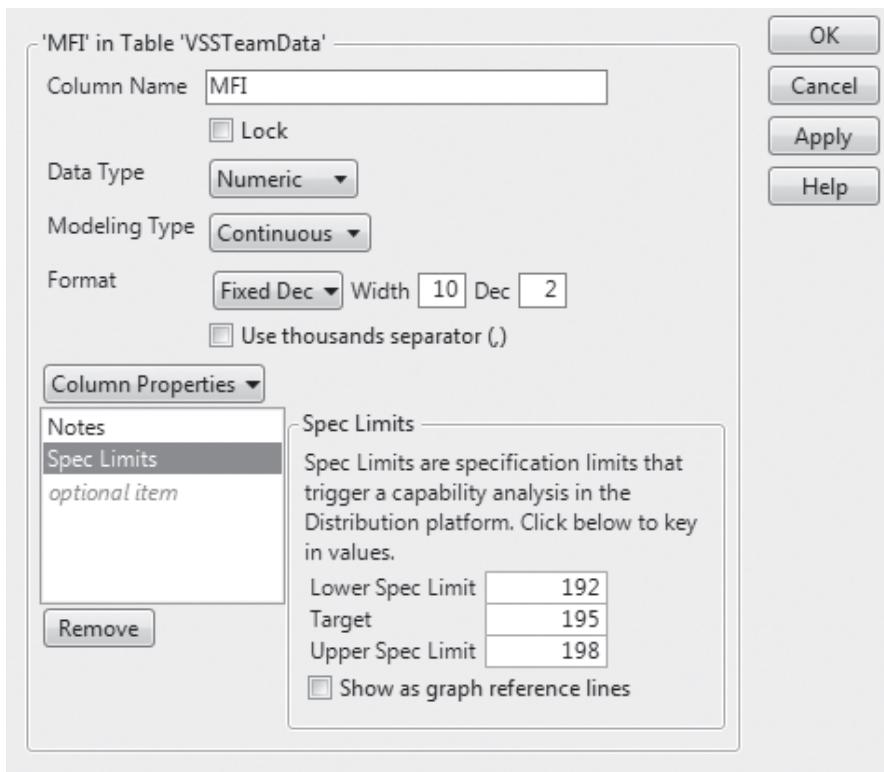


Exhibit 9.51 Spec Limits Column Property Panel for MFI

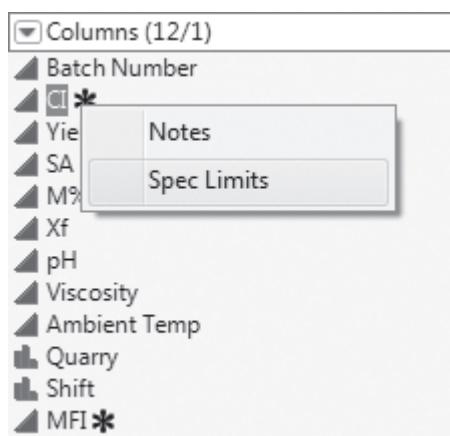


Exhibit 9.52 Asterisk Showing Spec Limits Column Property

Plan for Modeling

Now you are ready to build your model. You have seen evidence of nonlinearity between *CI* and *Xf* and between *MFI* and *SA*. How can you build this into a model?

You decide to get some advice on how to proceed from your mentor, Tom. Over lunch together, Tom suggests the following:

- Build a model for each of *MFI* and *CI* as responses using **Fit Model**.
- In the initial model, include response surface terms for all effects. Quadratic terms for the nominal effects do not make sense and will not be added. However, interaction terms involving the nominal terms will be added. Also include cubic terms in *Xf* and *SA*, based on the nonlinearity suggested by your bivariate analysis.
- Construct the models using the **Stepwise** personality. Tom shows you a quick example.
- Construct models for *MFI* and *CI* using the Hot Xs that you have identified.
- Use the **Profiler** to simultaneously optimize these models, thereby obtaining settings of the Hot Xs that optimize both *MFI* and *CI*.
- Quantify the anticipated variability in both responses, based on the likely variation exhibited by the Hot Xs in practice.

Tom's plan makes sense to you, and you proceed to implement it.

Building the Model

Your first step is to define your initial model. It will contain main effects, two-way interactions, quadratic effects for all continuous variables, and cubic terms in *Xf* and *SA*.



To fit your initial model (**VSSTeamData.jmp**, script for completed Model Specification window is **Model—Initial**):

1. Select **Analyze > Fit Model**.
2. Enter *MFI* and *CI* as **Y**.
This will produce a fit for each response in a single report.
3. In the **Select Columns** list, select the columns from *SA* to *Shift*.

4. From the **Macros** menu, select **Response Surface**.

This selection adds all main effect, two-way interactions, and all quadratic terms in the continuous variables. In the Construct Model Effects list, the notation **& RS** follows each continuous main effect, indicating that response surface terms for that effect have been added. There are 70 effects in all.

5. Select **SA** and **Xf** in the **Select Columns** list.
 6. Enter **3** in the **Degree** text box beneath the **Macros** button.
 7. From the **Macros** menu, select **Polynomial to Degree**.
This selection adds cubic terms in **Xf** and **SA**.
 8. From the **Personality** menu, select **Stepwise**.
 9. Click **Run**.
-

Stepwise Variable Selection

The **Stepwise** report appears (Exhibit 9.53), showing two main outline nodes, one entitled Stepwise Fit for **MFI** and the other entitled Stepwise Fit for **CI** (the **Current Estimates** panels have been minimized). The **Stepwise Regression Control** panel in each of these reports gives you control over how stepwise selection is performed. You decide to accept the JMP default settings, which specify a **Minimum BIC** stopping rule, the **Forward** direction, and the **Combine** rule that combines effects in determining their significance.

The *Bayesian Information Criterion*, or BIC, is a measure of model fit based on the likelihood function. It includes a penalty for the number of parameters in the model. A lower value of BIC indicates a better model. For details on the BIC and other selections, see www.jmp.com/support/help/Fitting_Linear_Models.shtml.



To use **Stepwise** to reduce your initial models for **MFI** and **CI** (**VSSTeamData.jmp**, script is **Stepwise Models**):

1. In the **Stepwise Regression Control** panel for **MFI**, click **Go**.
Model terms are selected in accordance with the control panel settings.
 2. Click **Make Model**.
A Model Specification window for **MFI** appears.
 3. In the **Stepwise Regression Control** panel for **CI**, click **Go**.
 4. Click **Make Model**.
A Model Specification window for **CI** appears.
-

Stepwise Fit for MFI

Stepwise Regression Control

SSE	DFE	RMSE	RSquare	RSquare Adj	Cp	p	AICc	BIC
464.4471	104	2.1132528	0.0000	0.0000	1132.3515	1	458.2179	463.4082

Current Estimates

Step History

Step	Parameter	Action	"Sig Prob"	Seq SS	RSquare	Cp	p	AICc	BIC
------	-----------	--------	------------	--------	---------	----	---	------	-----

Stepwise Fit for CI

Stepwise Regression Control

SSE	DFE	RMSE	RSquare	RSquare Adj	Cp	p	AICc	BIC
26140.89	104	15.85417	0.0000	0.0000	502.96672	1	881.4108	886.6011

Current Estimates

Step History

Step	Parameter	Action	"Sig Prob"	Seq SS	RSquare	Cp	p	AICc	BIC
------	-----------	--------	------------	--------	---------	----	---	------	-----

Exhibit 9.53 Stepwise Fit Window with Current Estimates Outlines Closed

Model Specification windows for *MFI* and *CI* are shown in Exhibit 9.54.

The Hot Xs for *MFI* are *SA*, *M%*, and *Xf*. The only Hot X for *CI* is *Xf*. You observe that *SA* and its quadratic and cubic effects appear in the model for *MFI* and that *Xf* and its quadratic and cubic effects appear in the model for *CI*.

Checking and Revising Stepwise Models

Next, you run each model to check the fit, starting with the model for *MFI*. In the Model Specification window for *MFI*, you click **Run**. Exhibit 9.55 shows the **Effect Summary** report and the **Actual by Predicted Plot**.

In the **Actual by Predicted** plot, you see a point that might be considered an outlier. Given its location, though, it doesn't appear to have high influence on the model fit. You don't see any other issues with the **Actual by Predicted Plot**, so you proceed to look at the effect tests.

(a)

(b)

Exhibit 9.54 Models Obtained Using Stepwise

The Stepwise Stopping Rule, Minimum BIC, can result in a model containing terms that are not *active* or significant. At this point, you reduce the model by removing terms that are not significant.

In the **Effect Summary** report, you notice that the interaction effect $M\% * Xf$ is not significant. Click on $M\% * Xf$ in the **Source** list to select it, as shown in

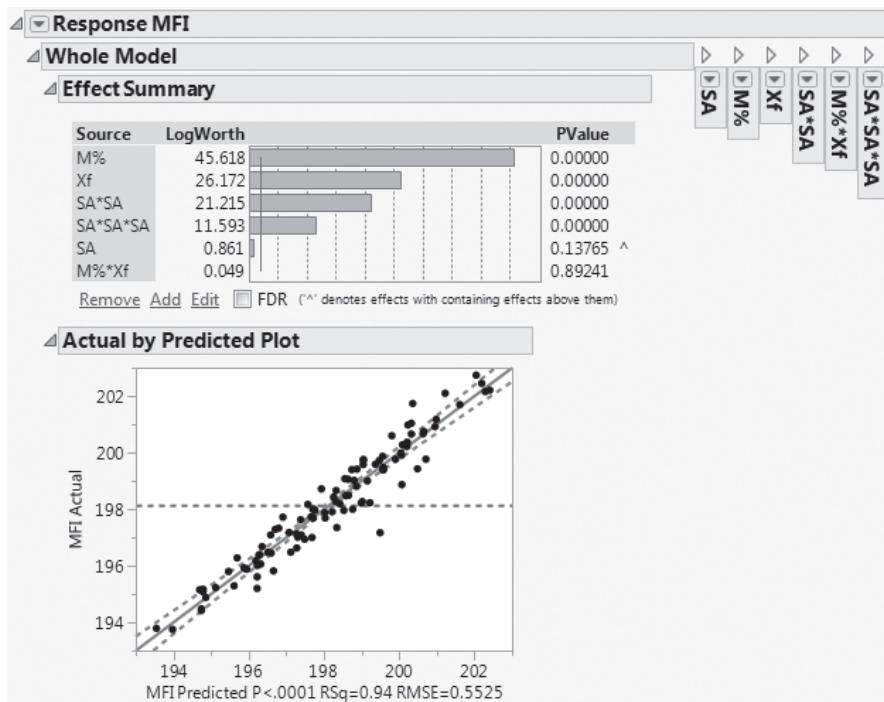
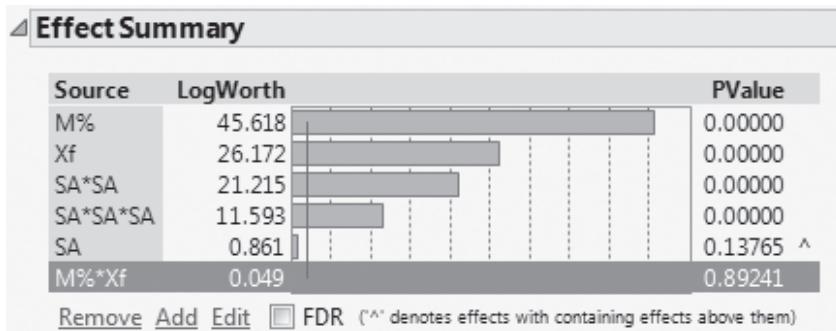
Exhibit 9.55 Partial Model Fit Report for *MFI*Exhibit 9.56 Effect Summary Panel for *MFI*

Exhibit 9.56. Then click **Remove**. This removes the effect from the model and updates all the reports in the window. Now all remaining effects are significant.

You want to save the model you have constructed to the data table as documentation for your analysis. To do this, from the **Response MFI** report's red triangle menu, select **Script > Relaunch Analysis**. Then save the model script to the data table. (The script for the saved model is **Model for MFI**.)

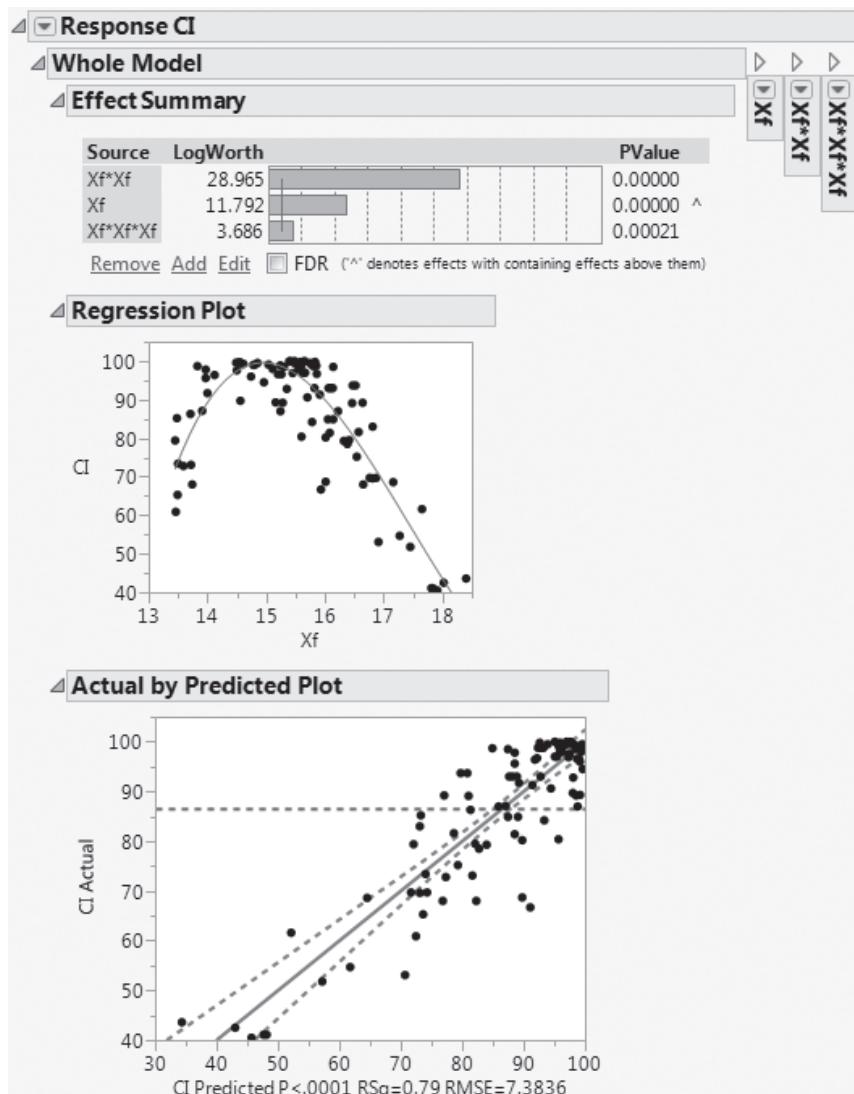


Exhibit 9.57 Partial Model Fit Report for *CI*

Next, run the **Stepwise** model for *CI*. Exhibit 9.57 shows the **Effect Summary** report and the **Actual by Predicted Plot**.

Based on the **Actual by Predicted** plot, the model seems to fit adequately. Because there is only one variable, namely *Xf*, in the model, JMP provides a **Regression Plot**. This shows how the model fits *CI* as a function of *Xf* and gives you an additional view of how the model fits. You see that the nonlinearity seems to be well modeled.

The **Effect Summary** report shows that all three effects are significant. The script for this model is saved to the data table as **Model for CI**.

All Possible Models

This seemed easy. You realize that your initial model has a large number of terms (70, to be exact, not counting the intercept). With that many terms, there are 2^{70} possible models (with intercept). This is a huge number, on the order of 10^{21} . Stepwise takes a selected path through the space of possible models that is based on p-values. This can be limiting.

You're tempted to do a quick check of your newfound models against other models that could be considered, using the **All Possible Models** option. You know that you will have to limit the search somehow because the total number of all possible models is daunting.



To obtain fits for a subset of all possible models for each of *MFI* and *CI* (**VSSTeamData.jmp**, scripts for sorted data tables are **All Possible Models for MFI** and **All Possible Models for CI**):

1. From the **Stepwise Fit** red triangle menu, select **All Possible Models**.
2. Enter **6** as the **Maximum number of terms in a model**.
Asking for more terms results in too many models for JMP to process.
3. Check the box next to **Restrict to models where interactions imply lower effects (Heredity Restriction)**.
This selection enforces strong effect heredity.
4. Click **OK**.
5. Click **OK** in the JMP Alert that appears.
6. Click **Continue** in the next JMP Alert that asks if you want to fit 10 million models.
Be patient as the calculation may take a while.
7. When the **All Possible Models** report appears, right-click in the report and select **Sort by Column**.
8. In the **Select Columns** window, select *B/C*, click **Ascending**, and then click **OK**.
This sorts the report by BIC in ascending order.
9. Right-click in the **All Possible Models** report and select **Make into Data Table**.

For *MFI*, the model with the smallest BIC, in row 1, is precisely the five-term model that you derived from the **Stepwise** model, once you removed the insignificant *M%*Xf* terms. For *CI*, the model with the smallest BIC, in row 1, is precisely the model selected by **Stepwise**.

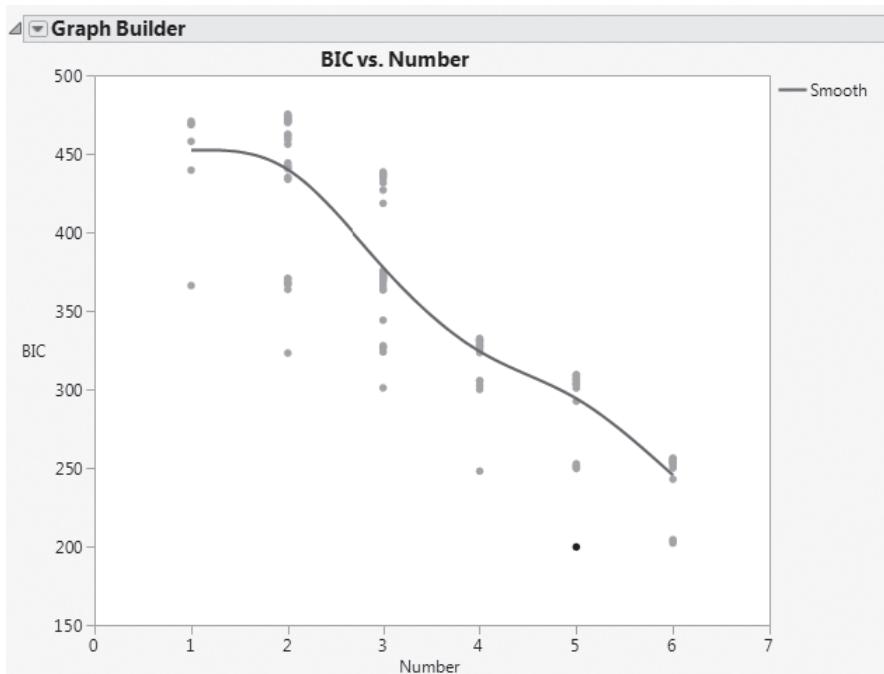


Exhibit 9.58 Plot of BIC versus Number of Terms for MFI

You can view the model information in the data tables produced by **All Possible Models** graphically by constructing a **Graph Builder** plot of *BIC* versus *Number*. The tables created by the scripts **All Possible Models for MFI** and **All Possible Models for Cl** contain **Graph Builder** scripts to construct these plots. Exhibit 9.58 shows the plot for *MFI*, with your final model selected.

Your **All Possible Models** analysis confirms that the terms that you included in your models for *MFI* and *Cl* are appropriate.

Saving the Prediction Formulas

Your plan is to obtain a prediction equation for each response and then to use the profiler to conduct multiple optimization.



To save the prediction formula for *MFI* to the data table (**VSSTeamData.jmp**), script is **Pred Formula for MFI**):

1. Run the script **Model for MFI**.
2. Click **Run**.
3. From the red triangle menu next to Response MFI, select **Save Columns > Prediction Formula**.

A column called *Pred Formula MFI* is added to the data table. To view the formula, right-click in the column header area and select **Formula**. This opens a formula editor box that shows the prediction equation.

Alternatively, in the **Columns** panel, the *Pred Formula MFI* column is now listed. To the right of the column name, you see a plus sign. Click on the plus sign to open the formula editor window displaying the formula.

Save the prediction formula for *CI* in the same fashion. The saved script is **Pred Formula for CI**.

REVISING KNOWLEDGE

At this point, you have constructed models for *MFI* and *CI* and you have identified the relevant Hot Xs. You and your team are ready to proceed to the Revise Knowledge step of the Visual Six Sigma Data Analysis Process. You will identify optimal settings for the Hot Xs, evaluate process behavior relative to variation in the Hot Xs using simulation, and run confirmatory trials to verify improvement.

Using the Profiler for Multiple Optimization

With both prediction formulas saved to the data table, you are ready to find settings for the Xs that will simultaneously optimize *MFI* and *CI*. Use **Graph > Profiler** to conduct multiple optimization.



Construct the **Prediction Profiler** plot in Exhibit 9.59 as follows (**VSSTeamData.jmp**, script is **Profiler**):

1. Select **Graph > Profiler**.
2. Select *Pred Formula MFI* and *Pred Formula CI* and enter these as **Y, Prediction Formula**.
3. Click **OK**.

From your training, you recall that each prediction formula defines a response surface. The **Profiler** shows cross-sections, called *traces*, of both response surfaces, with the first row of panels corresponding to *MFI* and the second row corresponding to *CI*. The cross-sections are given for the designated values of *SA*, *M%*, and *Xf*. You can drag the vertical red dotted lines using the

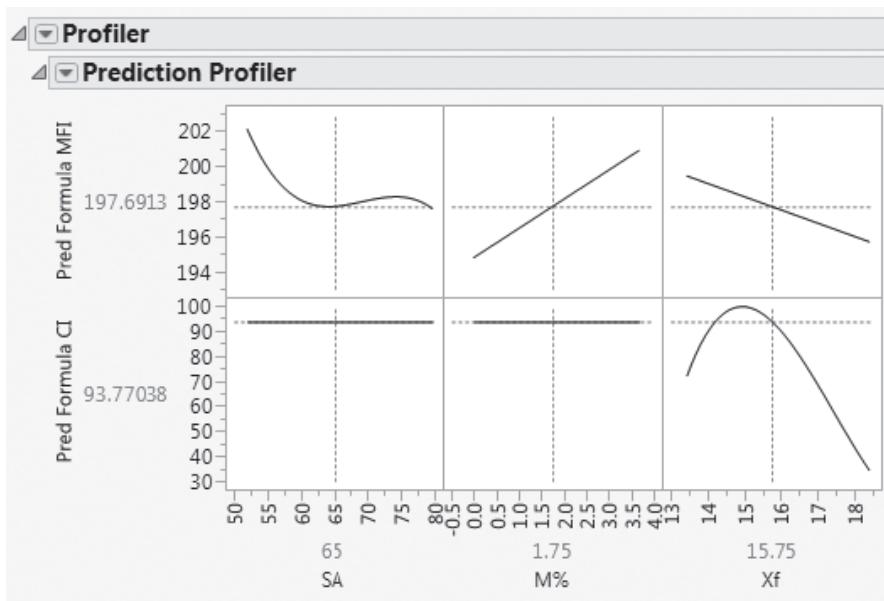


Exhibit 9.59 Initial **Profiler** Report

arrow tool to see how the two predicted responses change at various factor settings.

Consider the first row of plots, which relate to the prediction model for *MFI*. The three continuous factors *SA*, *M%*, and *Xf* define a response surface in four dimensions: There are three factor values, and these result in one predicted value. Given the settings in Exhibit 9.59, the **Profiler** is telling us, for example, that when *M%* = 1.75 and *Xf* = 15.75, the predicted *MFI* values, for various settings of *SA*, are those given in the panel above *SA*. In other words, for specific settings of any two of the continuous factors, the profiler gives the cross-section of the response surface for the third continuous factor.

Consider the second row of plots, corresponding to *Pred Formula CI*. The traces for *SA* and *M%* are perfectly horizontal. This is because *SA* and *M%* do not appear in the prediction formula for *CI*. The plot in the second row in the cell above *Xf* shows predicted values of *CI* for various values of *Xf*.

Understanding the Profiler Traces

It is getting late in the day, and the idea of visualizing these traces by plotting a response surface sounds like a wonderful diversion. You suspect that this will help you explain to your team members what these traces represent. You are

especially interested in visualizing the interesting trace for *MFI* above *SA* in the **Prediction Profiler** in Exhibit 9.59.



To visualize the trace for *SA* and to construct the plot in Exhibit 9.60 (**VSSTeamData.jmp**, script is **Surface Plot—SA and Xf**):

1. Select **Graph > Surface Plot**.
2. Enter *Pred Formula MFI*, *SA*, *M%*, and *Xf* as **Columns**.
3. Click **OK**.
A surface plot appears.
4. In the **Independent Variables** panel, select *SA* as **X** and *Xf* as **Y**.
5. Set *M%* to **1.75**, its value in the **Prediction Profiler**.
6. Set *Xf* to **15.75**, its value in the **Prediction Profiler**.
7. Check the box for **Grid** next to *Xf*.
8. In the **Dependent Variables** panel, click the color box to the right of *Pred Formula MFI*.
9. In the menu that appears, under **Fill Type**, select **Continuous Gradients**.
This applies a gradient coloring to *Pred Formula MFI*.
10. Click **OK**.
11. Click and drag your cursor in the plot to rotate it.

The vertical grid cuts the surface in a curved shape. This is precisely the trace shown for *SA* in the profiler in Exhibit 9.59 when *Xf* = 15.75 and *M%* = 1.75.

This exercise in three dimensions provides some intuition about the **Profiler** traces. When you change the settings of factors in the **Prediction Profiler**, the traces correspond to cross-sections of the prediction equation at the new settings.

Back to Simultaneous Optimization

“Well,” you think, “that was fun! I need to show that to the team members later on. They will love it.” Now you proceed to your next task, which is to find optimal settings for the Hot Xs. To do this, you return to the profiler. But, by now, you have lost that window. So you simply double-click on one of your reports, select **Window > Close All Reports**, and run the **Profiler** script that is saved to the data table.

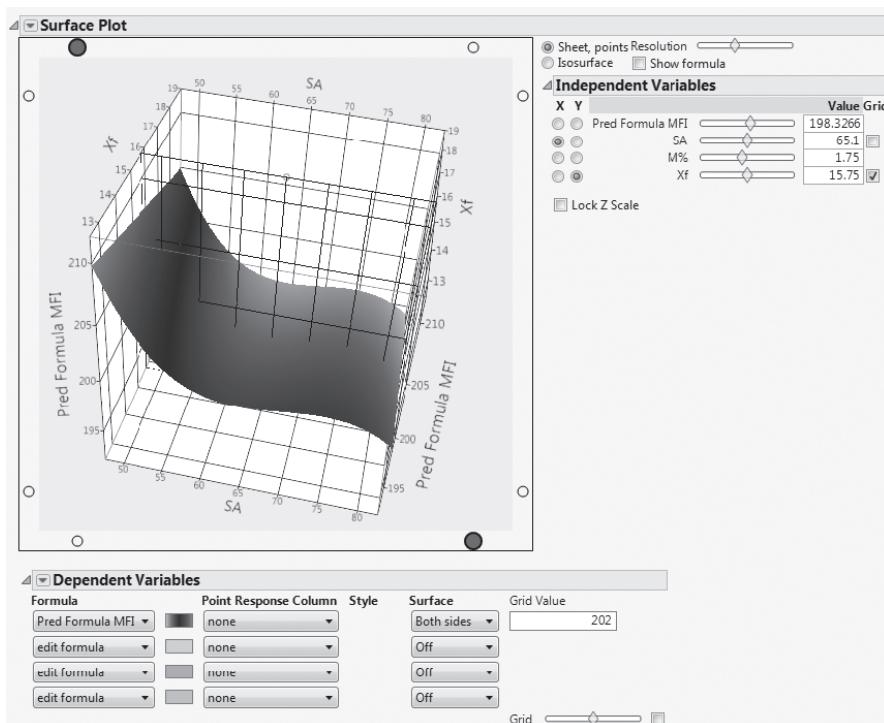


Exhibit 9.60 Surface Plot for Pred Formula MFI

To find optimal settings, you use the JMP desirability functions. Click on the red triangle next to **Prediction Profiler** and choose **Desirability Functions**. The saved script is **Profiler—Desirability Functions**.

This adds a row and a column of cells to the profiler. Two cells are added at the right, one for each of the responses (Exhibit 9.61). These show the desirability of various values of the responses, based on the specification limits that you entered into **Column Info** for *MFI* and *CI* earlier (see the section “Dealing with the Preliminaries”). For example, values of *MFI* near 195 have desirability near 1 while values above 198 and below 192 have desirabilities near 0. Similarly, large values of *CI* have high desirability.

To see the description of the desirability function for a given response, double-click in the desirability function panel for that response. For *MFI*, you obtain the **Response Goal** window shown in Exhibit 9.62. After reviewing the entries, click **Cancel** to exit with no changes.

A new row of cells, called **Desirability**, is added to the **Profiler** (Exhibit 9.61). This row gives traces for the joint desirability function, which, like the predicted surface, is a function of *SA*, *M%*, and *Xf*. For example, if *SA* = 65 and *M%* = 1.75, then low and very high values of *Xf* are not very desirable. Values of *Xf* around 16.3 are more desirable.

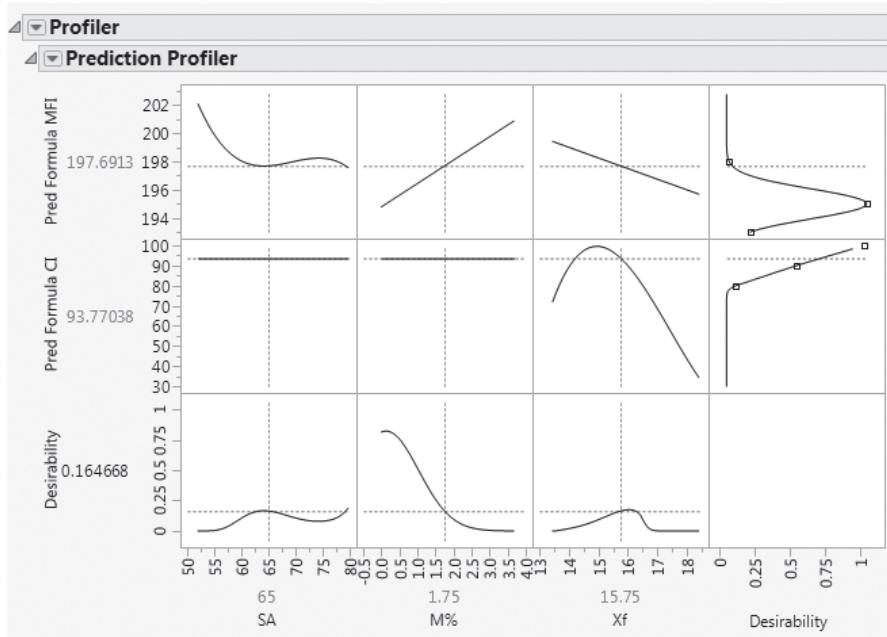


Exhibit 9.61 Profiler Showing Desirability Functions

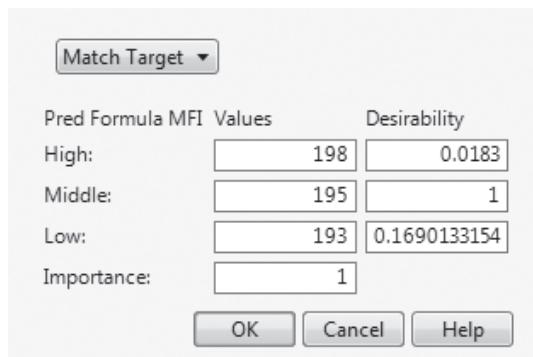


Exhibit 9.62 Response Goal Dialog for MFI

The goal is to find settings of the factors that maximize the desirability value. To find such settings, you select **Maximize Desirability** from the red triangle menu next to **Prediction Profiler**. The saved script is **Profiler—Desirability Maximized**.

You obtain the settings shown in Exhibit 9.63. The plot indicates that optimal settings for the three continuous Hot Xs are $SA = 79.7$, $M\% = 0$, and $Xf = 15.07$. At these settings, the predicted value of MFI is 195.1857 and of CI is 99.4645. Both predicted values are well within their specification limit ranges.

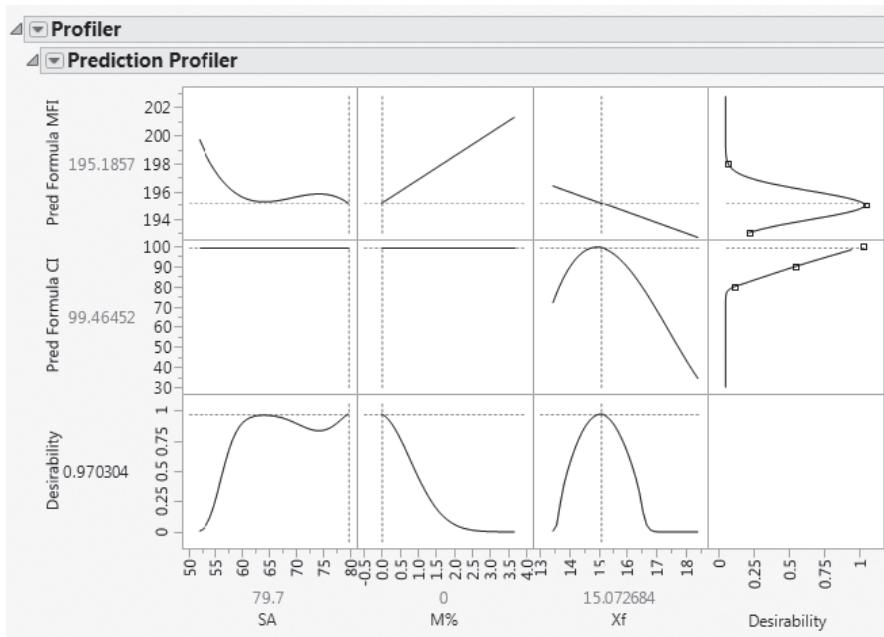


Exhibit 9.63 Profiler with Desirability Functions Maximized

Note that, when you maximize desirability, you may not obtain these exact settings. This is because, in most cases, settings that produce the same maximum value are not unique.

The fact that the optimal setting for *M%*, the viscosity modifier, is 0 indicates that adding the modifier actually degrades *MFI* (*M%* is not involved in the model for *CI*). This surprises your team members. It was always believed that the modifier helped the melt flow index. The team members talk with engineers about this, and it is decided that perhaps the modifier is not required—at least they are interested in testing a few batches without it.

The plot also indicates that, for the optimal settings of *SA* and *M%*, if *Xf* drifts to the 16.0 or higher, then desirability drops dramatically. By dragging the vertical dashed line for *Xf* to 16.5 and higher while looking at the *Pred Formula CI* trace for *Xf*, you see that *Pred Formula CI* drops below 80 when *Xf* exceeds 16.55 or so. *CI* also drops below 80 if *Xf* drifts to 13.65 or lower. You conclude that *CI* is very sensitive to variation in *Xf* at the optimal settings of *M%* and *SA*.

Assessing Sensitivity to Predictor Settings

In your training, you learned about a sensitivity indicator. The sensitivity indicator helps you assess how sensitive the predicted value is to factor settings.

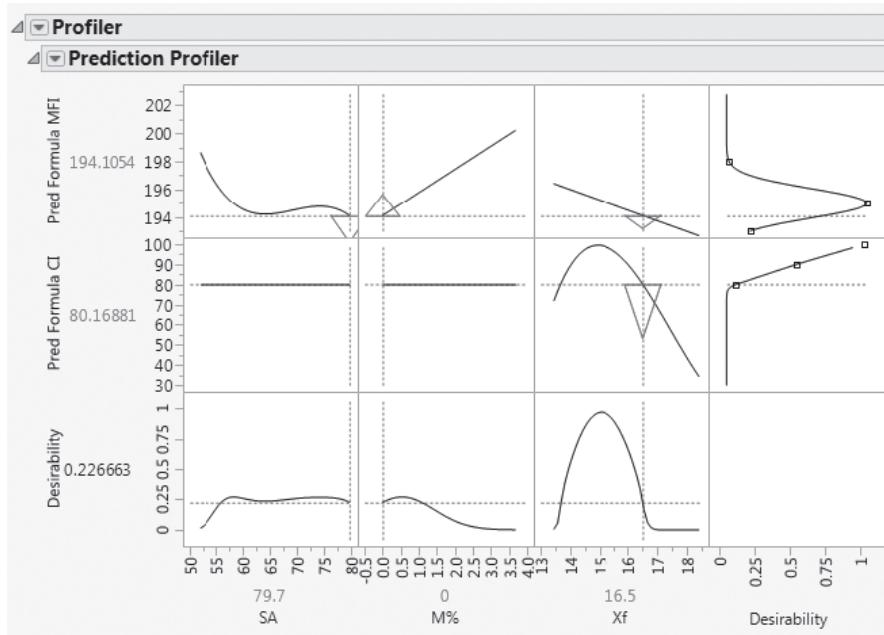


Exhibit 9.64 Profiler with **Sensitivity Indicators** and Xf near 16.5

From the **Prediction Profiler** red triangle, select **Sensitivity Indicator**. Hollow triangles are added to the prediction traces. Notice that no triangles are added to the traces for CI and SA or CI and $M\%$, since SA and $M\%$ are not involved in the prediction formula for CI .

You leave $M\%$ and SA at their optimal values and move the Xf setting to about 16.5 (Exhibit 9.64). The sensitivity indicator for CI in the Xf panel gets larger as you increase Xf to 16.5. The sensitivity indicator points in the direction of change for the response as the factor level increases and its size is proportional to the change. You conclude that maintaining tight control over Xf is critical. The saved script in **VSSTeamData.jmp** is **Profiler—Sensitivity Indicators**.

Now you want to see how sensitive the prediction is to SA at the optimized settings of $M\%$ and Xf . To reinstate the optimal settings, rerun the data table script **Profiler—Desirability Maximized** and select **Sensitivity Indicator** from the Profiler's red triangle menu.

Move the SA setting to its left. Note that the **Sensitivity Indicator** for MFI becomes very small at around 64.5, much smaller than it was at the optimal setting of 79.7, which was close to the upper bound of the experimental range. In fact, there is a range of values around 64.5 for which the desirability is almost as high as when SA is set to 79.7. See Exhibit 9.65 and compare the predicted values to those in Exhibit 9.59.

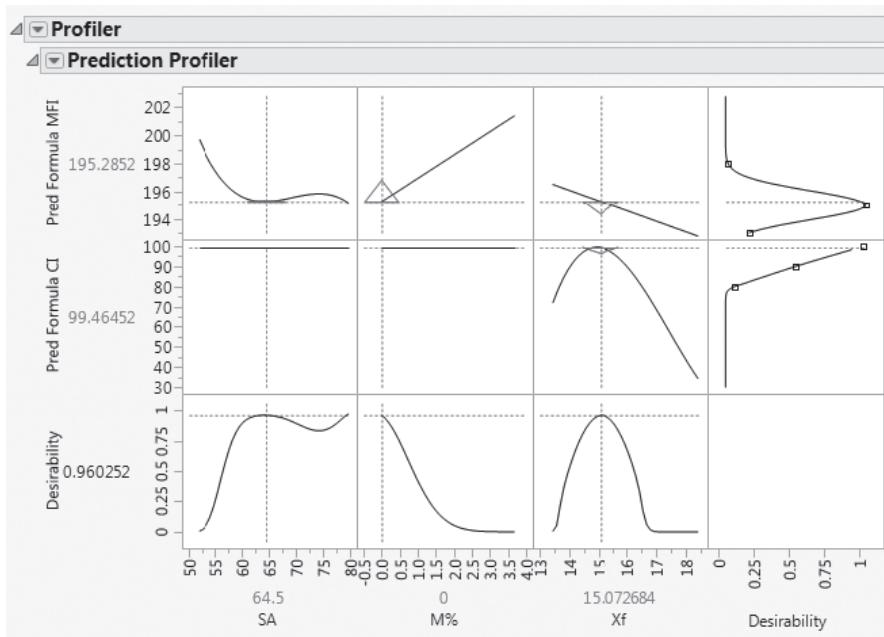


Exhibit 9.65 Profiler with **SA** Set to 64.5

Given the possibility of a more robust operating range for values of **SA** around 64.5, you want to consider using 64.5 as the optimal setting for **SA**. You do note, however, that the sensitivity indicator for **MFI** shows that **MFI** is highly sensitive to **SA** excursions to 60 and below.

Your sensitivity analysis allows you to reach two conclusions:

- Optimal settings that involve **SA** near 64.5 would result in a more robust process.
- The important sensitivities are those of **MFI** to low **SA** and **Cl** to high and low **Xf**.

Determining Optimal Process Settings

You realize that there are many combinations of settings of the factors that optimize both **MFI** and **Cl**. When you select **Maximize Desirability**, the profiler provides one such set of settings.

You have seen that setting **SA** to 64.5 results in essentially the same predicted values of **MFI** and **Cl** as did the settings provided when you maximized the desirability function. You decide to use 64.5 as the setting for **SA**. But you wonder if, for this setting of **SA**, there might be better settings for **M%** and **Xf**.

To find optimal settings for $M\%$ and Xf when SA is set to 64.5 (**VSSTeamData.jmp**, script is **Optimal Settings**):

1. **ALT**-click in the Desirability cell for SA .
2. Set the **Current Value** to **64.5**.
3. Check the **Lock Factor Setting** box.
4. Click **OK**.
5. Select **Maximize Desirability** from the **Prediction Profiler** red triangle menu.

Exhibit 9.66 shows that the optimal setting for $M\%$ is still 0 and that the optimal setting for Xf has changed only slightly, to 15.12.

Controlling SA and Xf

Your sensitivity analysis makes it clear that both SA and Xf must be controlled tightly around the optimal settings that you have derived. Your team members get started on developing an understanding of why these two factors vary and on proposing procedures that will ensure tighter control over them. They

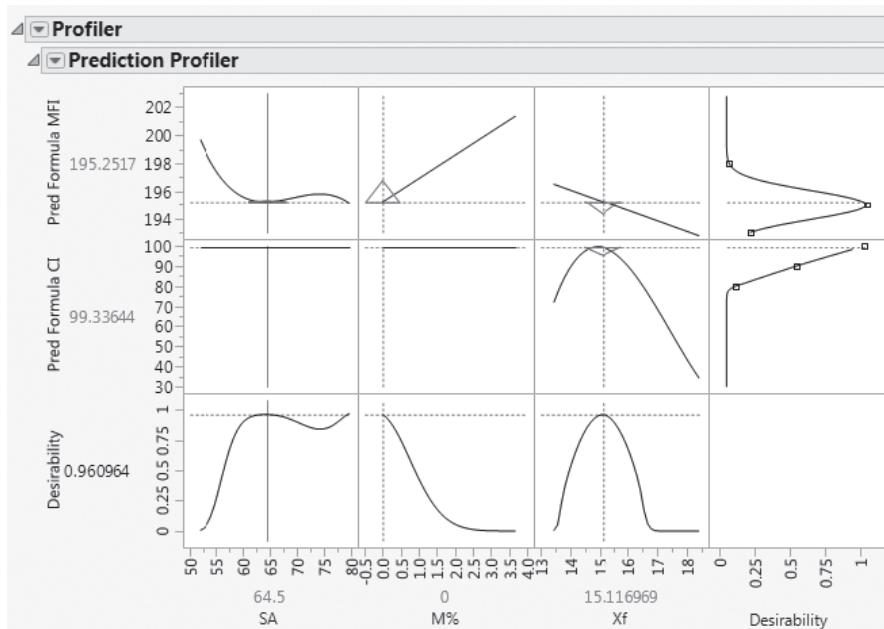


Exhibit 9.66 Optimal Settings with SA Locked at 64.5

propose a new control mechanism to monitor **SA** (amps for slurry tank stirrer) and to keep it on target at the optimal value of 64.5 and within narrow tolerances. They also develop procedures to ensure that filler and water will be added at regular intervals to ensure that **Xf** (percent of filler in the polymer) is kept at a constant level close to 15.12 percent.

The team tests these new controls and collects data from which you can estimate the variability in the **SA** and **Xf** settings in practice.

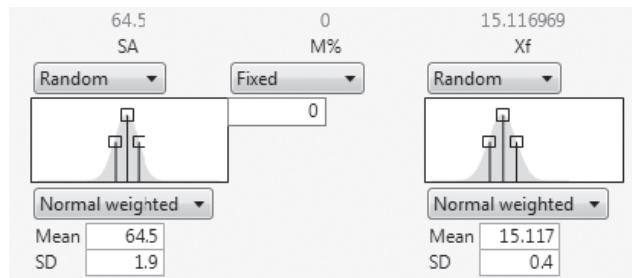
Simulating Process Outcomes

This last task takes a few weeks. But the resulting data indicate that when **SA** and **Xf** are held on target their actual settings do vary. Both factors appear to have approximately normal distributions with standard deviations of 1.9 and 0.4, respectively. Your plan is to use these estimates of variation, together with simulation, to investigate how variation in **SA** and **Xf** propagates to variation in **MFI** and **Cl**. In particular, you want to obtain reliable estimates of capability for both responses of interest.



To use the **Simulator** from the **Profiler** to see how variation propagates (**VSSTeamData.jmp**, script is **Profiler—Simulator**):

1. Rerun the script **Optimal Settings**.
2. From the red triangle menu next to **Prediction Profiler**, select **Simulator**.
This adds a **Simulator** outline to the report, as well as menus below the factor settings in the profiler. The dropdown menus enable you to specify distributions to use in simulating factor variability.
3. In the menu beneath **SA**, select **Random**.
4. In the dropdown menu beneath the normal plot that appears, select **Normal weighted**.
The **Normal Weighted** option samples from a normal distribution, but in a fashion that oversamples the tails of the distribution. This gives more precise estimates of the PPM defective rate than would otherwise be possible.
5. Enter **1.9** as the **SD** (see Exhibit 9.67).
Note that the mean is set by default at the optimal setting.
6. In the menu beneath **Xf**, select **Random**.
7. In the dropdown menu beneath the normal plot that appears, select **Normal weighted**.
8. Enter **0.4** as the **SD** (see Exhibit 9.67).
9. Click the **Simulate** button found at the right beneath the profiler cells.

Exhibit 9.67 Simulator with **Normal Weighted** Selected and **SD** Entered

Simulate					
Defect	Rate	PPM	Mean	SD	
Pred Formula MFI	5.33e-8	0.05331	195.303	0.3152	
Pred Formula CI	0.00041	413.497	97.789	2.39353	
All	0.00041	413.55			

Exhibit 9.68 Simulator Results

A **Defect** table appears below the **Simulate** button (Exhibit 9.68). Because the predicted values for *MFI* and *CI* are simulated, you will see different results than those shown here. If you run the script **Profiler—Simulator**, however, you will obtain the same results.

Right-click in the table and select **Columns > PPM**. When your team members see the overall **PPM** value of 413.55 as given in Exhibit 9.68, they give a cheer. You click on **Simulate** to run the simulation a few more times, and although there is some variation in the **PPM** values, the PPM values rarely exceed 500.

The **Defect** table indicates that the majority of defective values are coming from *CI*. Recall that *CI* is highly sensitive to *Xf* values in the region of the optimal settings. If the standard deviation for *Xf* could be reduced, this might further reduce the estimated **PPM** rate. Nonetheless, the team has very good news to report. Everyone is anxious to meet with Edward, the manufacturing director, and to test these predictions in the plant.

Confirming the Improvement

Edward is thrilled with the team's report and gives the go-ahead to run formal confirmation trials, measuring yield through to the molding plant. In the limited experience you have had to date with running the process at the optimal settings, you have only monitored the polymer plant yield.

	Batch Number	MFI	CI	Yield	SA	M%	Xf	Phase
1	4121	194.9	91.2	100.00%	62.5	0	15.0	Confirmation
2	4122	193.6	87.5	100.00%	62.9	0	14.9	Confirmation
3	4123	194.7	95.4	100.00%	63.0	0	14.6	Confirmation
4	4124	193.6	98.9	100.00%	62.2	0	14.9	Confirmation
5	4125	194.9	99.3	100.00%	62.7	0	15.1	Confirmation

All rows 5
Selected 0
Excluded 0
Hidden 0
Labelled 0

Exhibit 9.69 Confirmation Data from Five Batches

The team decides to perform tests on five batches using the target settings and controls for *SA* and *Xf*. The batches are followed through the molding plant so that, in line with the original definition, *Yield* is measured using both polymer plant and molding plant waste. The measurements for these five batches are given in the data table **ConfirmationData.jmp**, shown in Exhibit 9.69.

What a delight to see all *MFI* and *CI* values well within their specifications and to see 100 percent *Yield* on all five batches! You and your team members are ecstatic.

UTILIZING KNOWLEDGE

In spite of this very promising result, you remind your teammates to remain cautious. Even if you had made no process changes, history shows that a crisis is only expected about once every quarter. During the single crisis that occurred over the team's baseline period, *MFI* values were unusually high. So there is statistical evidence that keeping *MFI* within the specification limits will alleviate crisis periods. But it will be important to continue to monitor *Yield* to ensure that the new optimal settings and controls really do address the root cause of a crisis. As you gather additional data over a longer time span, you will have more assurance that keeping *MFI* and *CI* within their specification limits really does alleviate crises.

This is an interesting situation where, although success has been confirmed in the short term, the longer time horizon of the historical problem requires a heightened sense of vigilance during the control phase. So, although you are ready to see this project move to the *Utilize Knowledge* step of the Visual Six

Sigma Data Analysis Process, that step retains some aspect of providing confirmation that the problem has been solved.

With this in mind, you and your team report your results to Edward. He is very impressed by the outcome of the confirmation runs and is extremely pleased by your team's work. He appreciates the logic and discipline you have imposed throughout the knowledge-building process. It appears that because of your diligence, the crisis periods may finally be eliminated.

Edward approves the fitting of a new control mechanism for the fill process to monitor *SA* and to keep it on target and within tight control. He approves of the procedures proposed by your team to maintain *Xf* on target with little variation. These procedures are shared with technicians and operators in a formal training session. Specification limits are defined for *SA* and *Xf* based on the standard deviations used in the simulation study.

Once these procedures are implemented, production batches are monitored closely, to ensure that nothing goes amiss, for the first month of production. Data on *MFI*, *CI*, *Yield*, *SA*, *M%*, and *Xf* for the 79 batches produced that month are given in the data table **ControlPhaseData.jmp**.

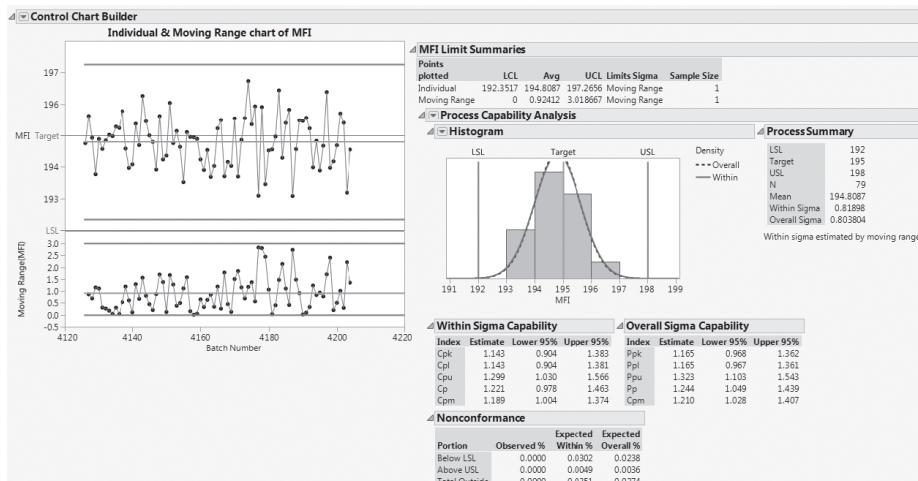
Verifying Process Stability

Open **ControlPhaseData.jmp**. The first thing that you notice is that all *Yield* values are 100 percent! But you are interested in the process measurements that drive *Yield*, namely, *MFI* and *CI*. The first thing that you check is whether *MFI* and *CI* are stable.

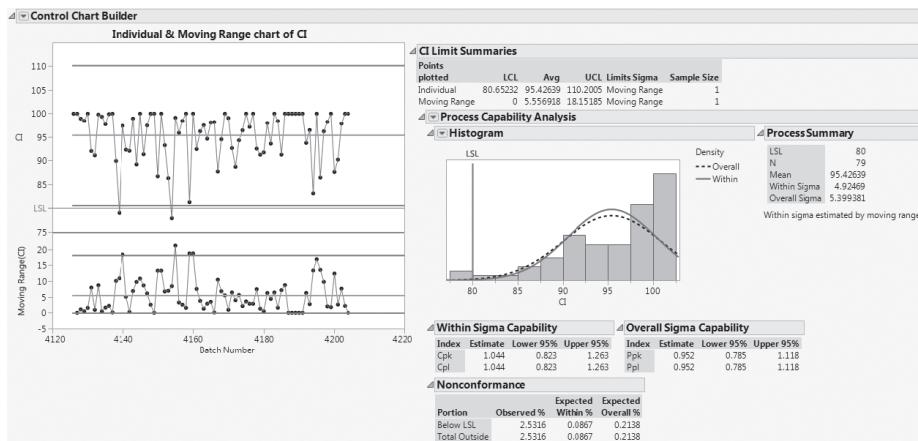
Use **Control Chart Builder** to construct Individuals and Moving Range charts for *MFI* and *CI* (script is **Control Charts**). Because specification limits have been entered in the columns for these two variables, JMP automatically computes normality-based capability analyses. The Individuals and Moving Range control charts are shown in Exhibit 9.70.

The *MFI* measurements are stable and fall well below the upper specification limit of 198. But the *CI* measurements appear unstable. However, note that their distribution is highly skewed, as we would expect. The individual measurement chart's control limits are valid for normally distributed data, and *CI* is not normally distributed. What to do?

Even in a case where the underlying data are highly non-normal, the means of samples are approximately normal, even for small samples or subgroups (this follows from the Central Limit Theorem). This means that an XBar chart can be used to assess process stability. So, for *CI*, you decide to construct an XBar and R chart, subgrouping the measurements into consecutive subgroups of size five.



(a)



(b)

Exhibit 9.70 Control Charts for Control Phase Data

To construct an XBar and R chart (**ControlPhaseData.jmp**, script is **XBar Chart for CI**):

1. Select **Analyze > Quality and Process > Control Chart Builder**.
2. Select **CI** and click the **New Y Chart** button beneath the **Select Columns** list.
3. From the red triangle menu next to Control Chart Builder, select **Set Sample Size**.
4. In the window that opens, a default sample size of 5 is specified. Click **OK**.

The resulting chart is shown in Exhibit 9.71. The absence of out-of-control points and unstable patterns is evidence that the *C/I* measurements have come from a stable process.

Now that you know that *MFI* and *C/I* are stable, you are interested in assessing their capability relative to meeting the specifications.

Estimating Process Capability

You return to the report obtained when you first constructed your Individuals and Moving Range charts using the data table **ControlPhaseData.jmp**. The script is **Control Charts**.

The Process Capability Analysis for *MFI* is repeated in Exhibit 9.72. The *MFI* distribution falls within the specification limits. Both short- and long-term capability indices are over 1.0: Cpk is 1.143, based on the moving range, and Ppk is 1.165. These values are close, as expected, because the process is stable.

You right-click in the Nonconformance report and select **Columns > Expected Overall PPM**. This selection adds a column that shows predicted parts per million out of spec. The calculation uses the standard deviation for all values, which is reasonable, since the process is stable. The estimated defective rate is 273.6 batches per million.

Since *C/I* is highly skewed, a capability determination based on a normal distribution is not meaningful and potentially misleading. In this case, the Cpk value, based on a normal fit, is 1.044. This seems optimistic.

A team member asks, "What do we do to assess capability?"

"Well," you reply, "let's try a nonparametric approach!"

For a left-skewed distribution, such as this one, it makes sense to use a method that does not specify a particular functional shape or form.



To obtain a capability analysis using a nonparametric fit (Exhibit 9.73), use **Distribution** (**ControlPhaseData.jmp**, script is **Capability C/I**):

1. Select **Analyze > Distribution**.
2. Enter *C/I* as **Y, Columns**.
3. Click **OK**.
4. From the red triangle menu next to *C/I*, select **Continuous Fit > Smooth Curve**.

The nonparametric fit is smooth but a bit bumpy. Yet, it describes the distribution quite well. The computed capability index is 0.775, and the PPM rate is 23,518, or about 2.35 percent. Based on the graph, your team is comfortable

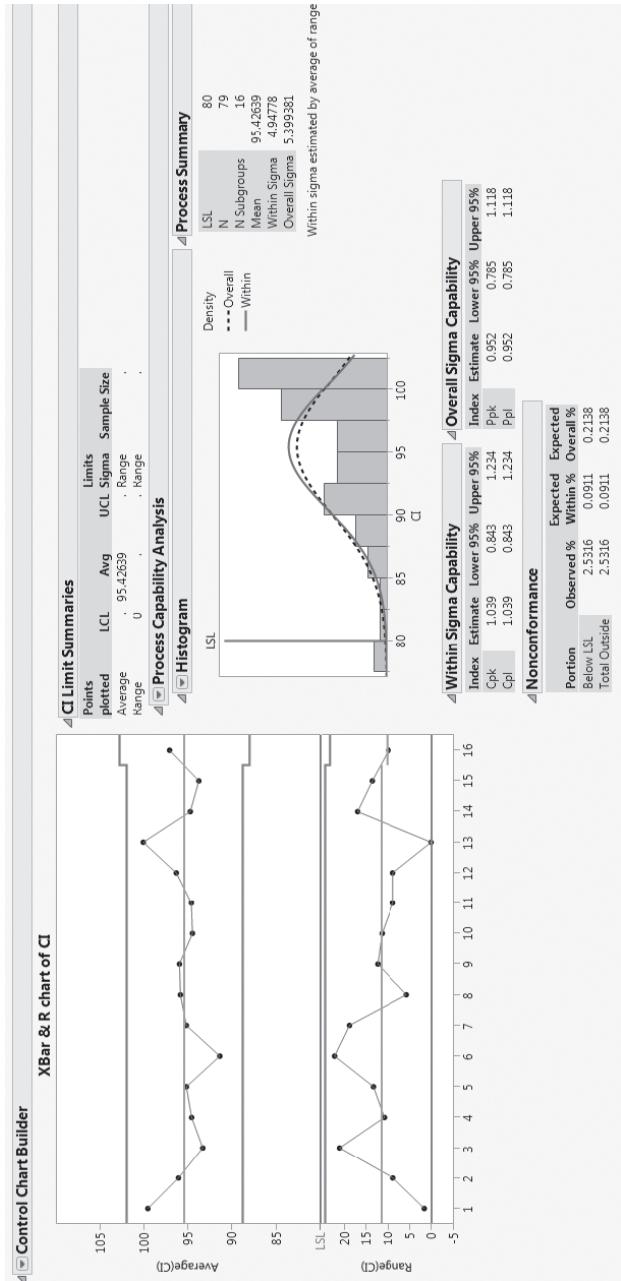


Exhibit 9.71 XBar and R Chart for Subgrouped CI Measurements

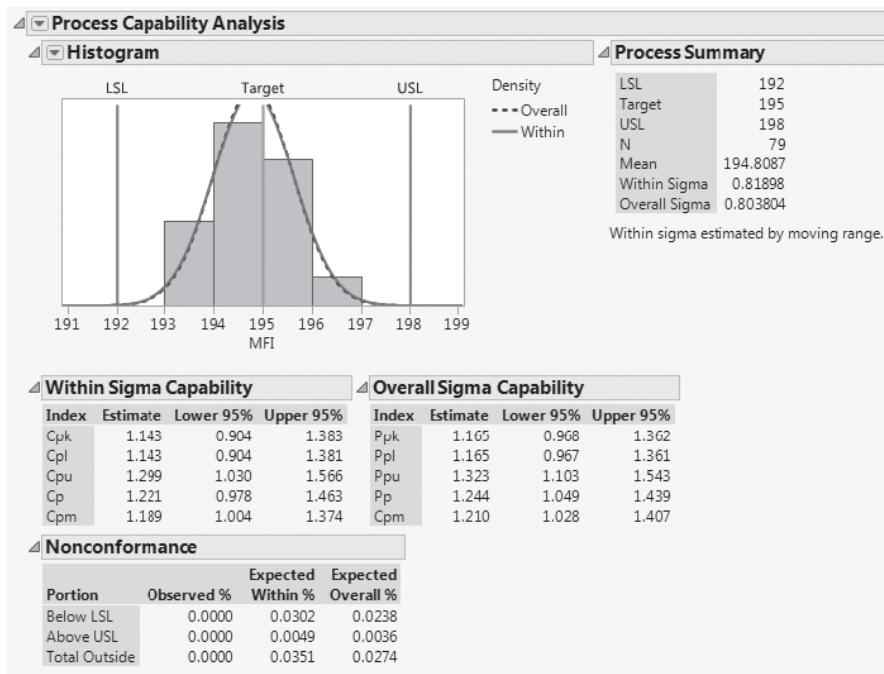


Exhibit 9.72 Process Capability Analysis for MFI

with this estimate of capability and PPM. The normality-based Cpk value of 1.044 was optimistic and incorrect.

It is also apparent to your team members that their simulation underestimated the PPM rates for both *MFI* and *CI*. Recall that the combined estimate of PPM was about 400. The actual data, especially for *CI*, shows a much higher estimated PPM rate. This may indicate that additional variability is coming from other sources.

The 2.35 percent out-of-specification rate for *CI* is clearly not acceptable. This is an area that will need further investigation. One of the team members suggests that the lower specification limit of 80 may in fact be too tight, given that all batches had a *Yield* of 100 percent. Certainly, if it is possible to loosen the specification without incurring any loss in quality, this is a worthwhile endeavor. You document the need for a resolution to this *CI* capability issue and put it on a list of future projects.

Tracking Improvement

To visualize progress over the life of the project, you want to concatenate the tables **VSSTeamData.jmp**, **ConfirmationData.jmp**, and **ControlPhase-Data.jmp** into a single data table. You want to be able to plot your data over

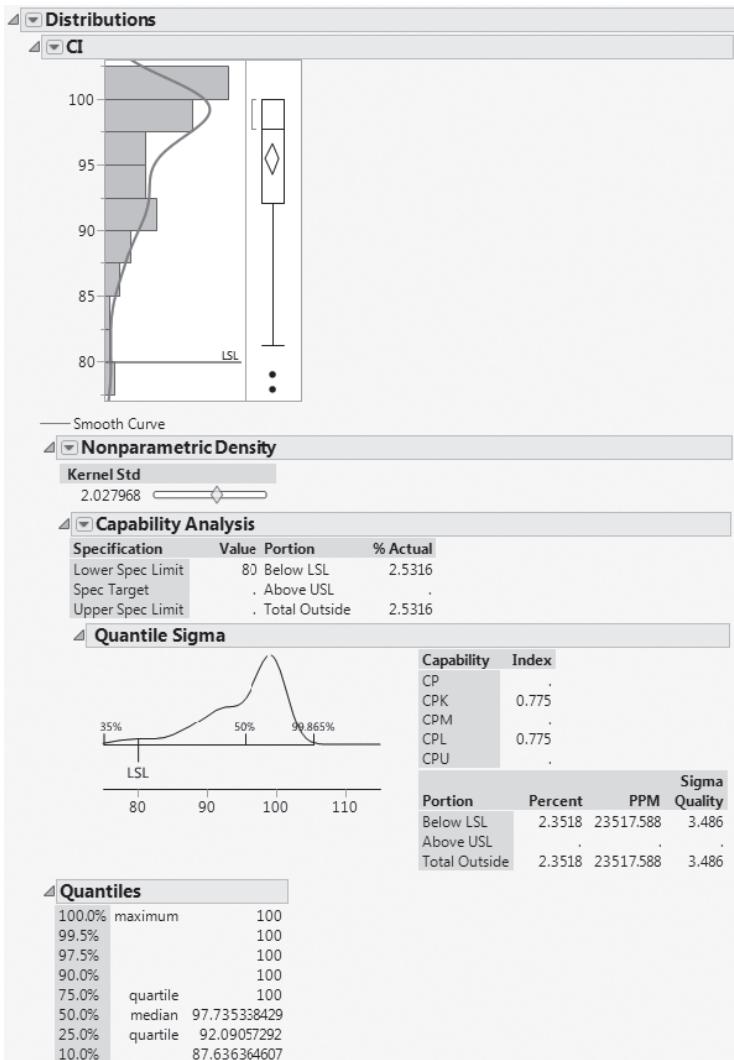


Exhibit 9.73 Smooth Curve Fit to CI

three phases: the phase consisting of the old process, the confirmation phase, and the control phase.

Before concatenating your tables, you engage in some housekeeping. Check your list of open windows under **Window**. If it shows a lot of open windows (and you're on a Windows machine), you can simply select **Windows > Close All** and reopen the three data tables. But the three data tables that you need may be on the dropdown list of open data. To close all but the data tables, select a report to make it active and then select **Window > Close All Reports**. Close the data tables that are not of interest. When you are finished, make sure that only the three data tables mentioned above are open.

You want to concatenate the three data tables, meaning that you want to append rows from one data table to another. You want your concatenated data

<input checked="" type="checkbox"/> VSSTeamData	<input checked="" type="checkbox"/> ConfirmationData ►	<input checked="" type="checkbox"/> ControlPhaseData ►
Graph Builder - Yield and CI Rows: <input checked="" type="checkbox"/> Distribution <input checked="" type="checkbox"/> Data Filter <input checked="" type="checkbox"/> Control Charts <input checked="" type="checkbox"/> Markers for Outliers <input checked="" type="checkbox"/> Graph Builder - Yield and MFI <input checked="" type="checkbox"/> Graph Builder - Yield and CI	<input checked="" type="checkbox"/> Distribution	<input checked="" type="checkbox"/> Control Charts <input checked="" type="checkbox"/> XBar Chart for CI <input checked="" type="checkbox"/> Capability CI
<input checked="" type="checkbox"/> Columns (12/0)	<input checked="" type="checkbox"/> Columns (8/0)	<input checked="" type="checkbox"/> Columns (8/0)
<input checked="" type="checkbox"/> Batch Number <input checked="" type="checkbox"/> MFI <input checked="" type="checkbox"/> CI <input checked="" type="checkbox"/> Yield <input checked="" type="checkbox"/> SA <input checked="" type="checkbox"/> M% <input checked="" type="checkbox"/> Xf <input checked="" type="checkbox"/> pH <input checked="" type="checkbox"/> Viscosity <input checked="" type="checkbox"/> Ambient Temp <input checked="" type="checkbox"/> Quarry <input checked="" type="checkbox"/> Shift	<input checked="" type="checkbox"/> Batch Number <input checked="" type="checkbox"/> MFI * <input checked="" type="checkbox"/> CI * <input checked="" type="checkbox"/> Yield <input checked="" type="checkbox"/> SA <input checked="" type="checkbox"/> M% <input checked="" type="checkbox"/> Xf <input checked="" type="checkbox"/> Phase	<input checked="" type="checkbox"/> Batch Number <input checked="" type="checkbox"/> MFI * <input checked="" type="checkbox"/> CI * <input checked="" type="checkbox"/> Yield <input checked="" type="checkbox"/> SA <input checked="" type="checkbox"/> M% <input checked="" type="checkbox"/> Xf <input checked="" type="checkbox"/> Phase

Exhibit 9.74 Column Panels for Three Data Tables

table to contain columns showing the basic information and a column for the Phase of the study.

So you start by examining the column structure in the three data tables (Exhibit 9.74). Note that **ConfirmationData.jmp** and **ControlPhase.jmp** have the same column structure, but **VSSTeamData.jmp** contains columns that the previous two data tables do not.

Follow the two sets of instructions below to concatenate the three tables. Alternatively, simply open the concatenated data table, called **VSSTeamFinal-Data.jmp**.



To adjust the column structure of the three data tables and to add a *Phase* column:

1. Click on **VSSTeamData.jmp** to make it active and save it as **TempData.jmp**.
The name indicates that you can delete it later on.
2. Select **Rows > Clear Row States**.
This ensures that no rows will be hidden and excluded in the concatenated table.
3. In the **Columns** panel of **TempData.jmp**, select all columns that follow *Xf*, right-click in the **Columns** panel, and select **Delete Columns**.

4. Double-click in the column header area to the right of the last column, *Xf*.
This creates a new column.
 5. Right-click in the column area and select **Column Info**.
 6. In the Column Info window, for **Column Name**, type *Phase* and click **OK**.
 7. In row 1 of the *Phase* column, type **Old Process** and then click away from this cell.
 8. Right-click back into the cell and select **Fill > Fill to end of table**.
This copies the text “Old Process” into all of the data table rows.
-

Now that your tables are prepared, you can concatenate them.



To concatenate the three tables:

1. Select **Tables > Concatenate**.
Since **TempData.jmp** is the active data table, it appears first in the list of *Data Tables to be Concatenated*.
 2. From the **Opened Data Table** list, select **ConfirmationData** and click **Add**.
 3. From the **Opened Data Table** list, select **ControlPhaseData** and click **Add**.
 4. In the **Output table name** text box, enter **MyFinalData.jmp**.
See Exhibit 9.75.
 5. Click **OK**.
-

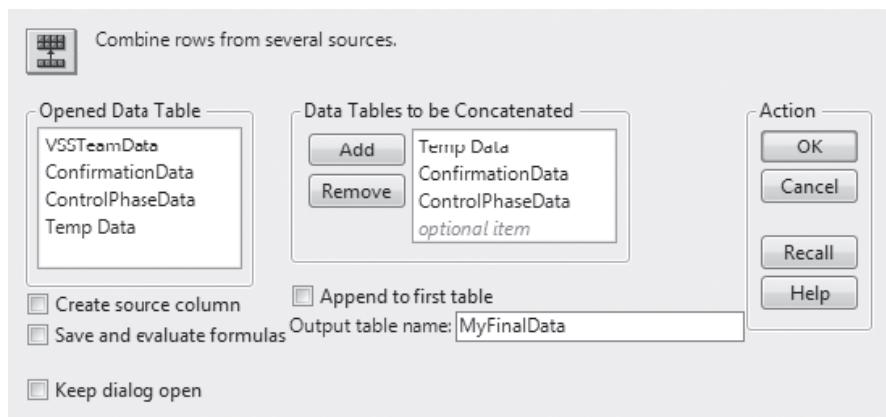


Exhibit 9.75 **Concatenate** Dialog for the Three Tables

The new table appears. You check to make sure it is what you intended. Close **TempData.jmp** without saving it: Select **File > Close** and select **No** when asked about saving changes.

Notice that all scripts have been carried over to the new table. Also, there is a new script called **Source**. This script tells you how the current data table was obtained and is useful for documentation. Some of the old scripts may not be meaningful since they were developed for portions of the data. However, you want to see control charts by *Phase*. You realize that the **Control Chart** and **XBar Chart for CI** scripts could be useful. To avoid confusion, you delete all the scripts other than **Source**, **Control Charts**, and **XBar Chart for CI**. (If you prefer, at this point, you may close your concatenated data table and open the saved data table called **VSSTeamFinalData.jmp**.)

In the data table **VSSTeamFinalData.jmp**, notice that *Phase* has the **Row Order Levels** property. Click on the asterisk to the right of *Phase* in the **Columns** panel to see this. This property ensures that the levels of *Phase* appear in plots in order of their occurrence in the data. When you plot control charts by *Phase*, the order of levels will be: Old Process, Confirmation, Control.



Obtain IR charts for *MFI*, *CI*, and *Yield* by *Phase* as follows (**VSSTeamFinalData.jmp**, script is **Control Charts by Phase**):

1. Run the script **Control Charts**.
This chart was carried over from **VSSTeamData.jmp**.
2. Select *Phase* from the **Select Columns** list and drag it to the **Phase** zone at the top of the template.
3. In turn, select *MFI*, *CI*, and *Yield* from the **Column Switcher** list at the far left of the report.

The *MFI* control chart, shown in Exhibit 9.76, shows how the measurements have changed over the three phases of the study. The *MFI* measurements are now less variable and within specifications.

Since the distribution for *CI* is skewed, the control limits for the Individuals Chart for *CI* are not appropriate. As earlier, you construct an XBar and R chart with subgroups of size 5.

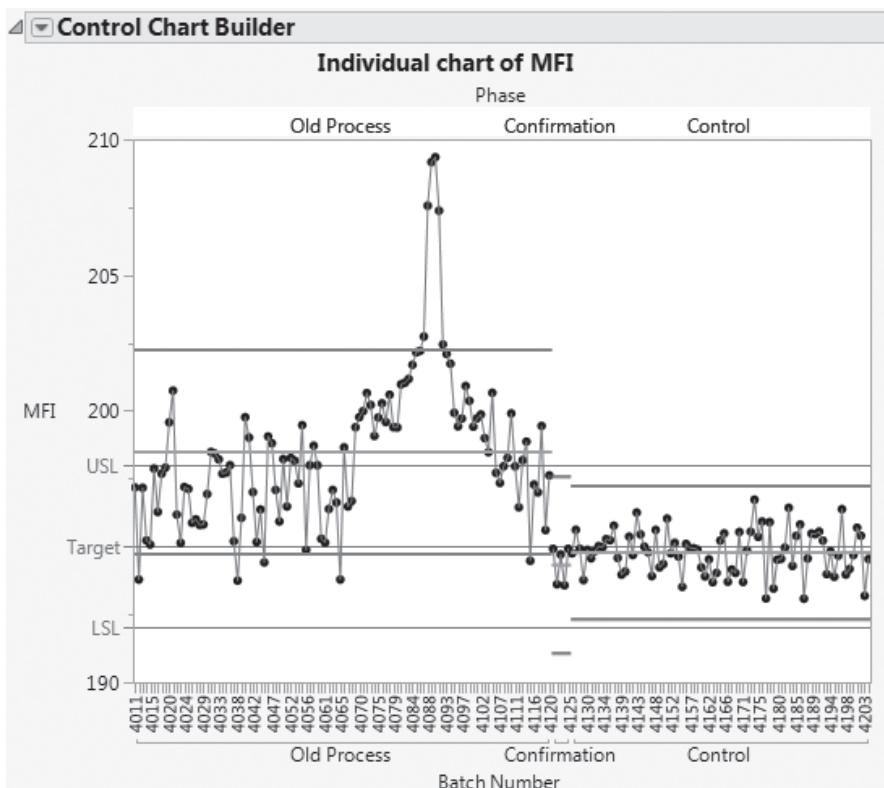


Exhibit 9.76 Phase Control Chart for *MFI*

To construct an XBar and R chart for *CI* by phase, shown in Exhibit 9.77 (the saved script is **XBar Chart for CI by Phase**):

1. Run the script **XBar Chart for CI**.
2. Select **Phase** in the Columns list and drag it to the **Phase** zone.

Starting with the confirmation phase, the *CI* process looks stable, both in terms of centering and spread. It is operating with much less variability and appears to be centered at about 95. Since this is an **XBar** chart, we cannot compare the control limits (based on averages) directly with the specification limits

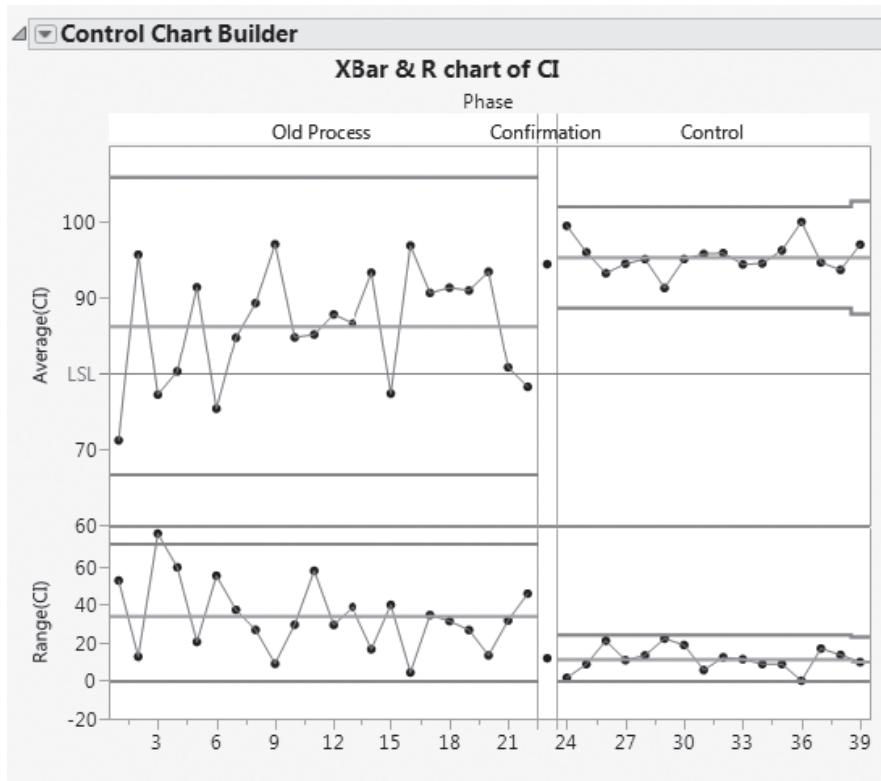


Exhibit 9.77 Phase XBar and R Chart for *CI*

(based on individual values). But that has already been done for the Control phase in the non-normal capability analysis (Exhibit 9.73).

CONCLUSION

The original project goal was to eliminate the periods of crisis with a goal of achieving a yield rate of 95 percent. One and a half years have now passed since completing this project, and although sometimes a very few batches do fall slightly below the lower limit for color index (CI), this issue is resolved by blending these low-color index batches with high-color index batches. To date, the molding plant has not rejected a single batch of white polymer since the changes have been implemented.

The savings from rejected batches alone are slightly over £750,000 per annum. Additionally, there are now no processing restrictions on the molding plant, resulting in an additional annualized margin of £2,100,000 per annum and a very happy customer. The actual costs incurred on the project were minimal.

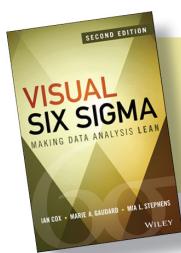
This was the project that convinced the skeptics, Edward included. It solved a major manufacturing problem with huge business benefits. In the final project close-out meeting, you summarized the key lessons as follows:

- Never trust a measurement system—it was only when these issues were resolved that it was possible to get meaningful results from the analysis.
- The DMAIC methodology in conjunction with the Visual Six Sigma Roadmap delivers high levels of process understanding in a simple, straightforward manner.
- Visual Six Sigma techniques enable technical and business users to apply the principles of statistical thinking to solve problems rapidly and effectively.
- Visual Six Sigma greatly facilitates the communication of results to a wide audience in a simple visual form.

NOTES

1. D. J. Wheeler, *EMP III: Evaluating the Measurement Process σ^2 Using Imperfect Data* (SPC Press, 2006).
2. Wheeler, *EMP III*.
3. Automotive Industry Action Group, *Measurement Systems Analysis Reference Manual*, 4th edition (Chrysler Group LLC, Ford Motor Company, General Motors Corporation, 2010).

From *Visual Six Sigma: Making Data Analysis Lean, Second Edition* by Ian Cox, Marie A. Gaudard, and Mia L. Stephens. Copyright © 2016, SAS Institute Inc., Cary, North Carolina, USA. ALL RIGHTS RESERVED.



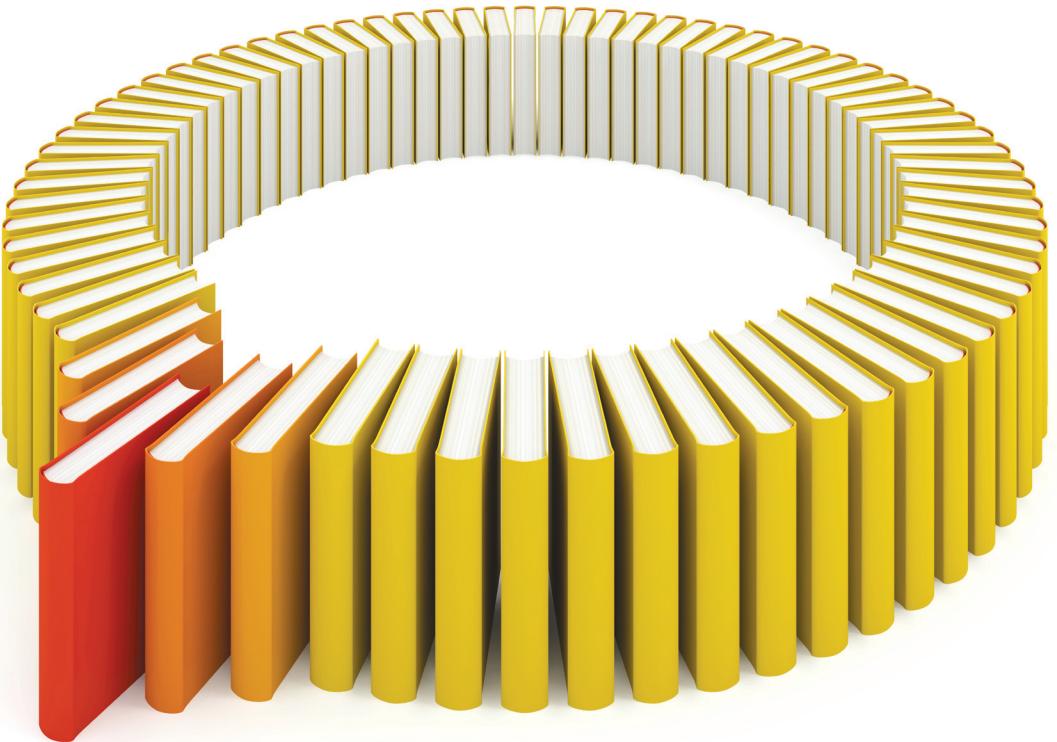
From *Visual Six Sigma: Making Data Analysis Lean, Second Edition*. Full book available for purchase [here](#).

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.



Gain Greater Insight into Your SAS® Software with SAS Books.

Discover all that you need on your journey to knowledge and empowerment.



support.sas.com/store/books
for additional books and resources.

sas
THE POWER TO KNOW®