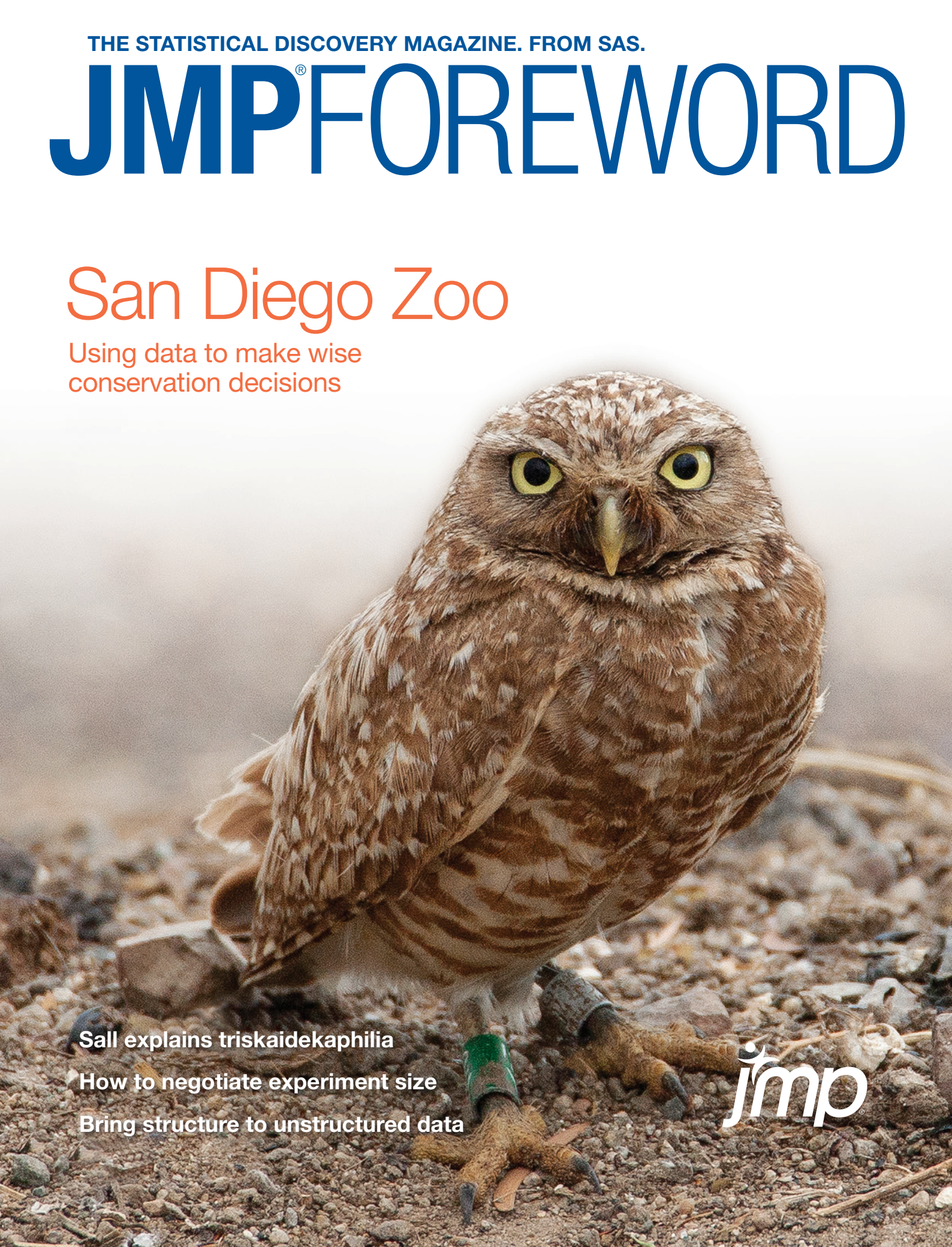
THE STATISTICAL DISCOVERY MAGAZINE. FROM SAS.

JMP® FOREWORD

San Diego Zoo

Using data to make wise conservation decisions



Sall explains triskaidekaphilia

How to negotiate experiment size

Bring structure to unstructured data

Not everyone gets to the summit

You may have heard: Discovery Summit is easier to get to than ever. The premier JMP conference series, which began in the US, has expanded into Europe, Japan and China. But wait: It's still not for everyone.

Discovery Summits are designed for the analytically inclined—those who want to exchange best practices in data exploration, and play with both new and tried-and-true statistical techniques.

These conferences put the best and brightest data explorers together. Why? Because together they innovate and leave inspired to spread the use of analytics.

Think you can keep up? If so, pick a Summit and reserve your seat.
jmp.com/discoverysummit

DISCOVERY SUMMIT

EXPLORING DATA
INSPIRING INNOVATION

jmp® Statistical Discovery.™ From SAS.

sas® THE POWER TO KNOW®

WHO'S WHO

EDITOR
Jessica Marquardt

FEATURES EDITOR
Meg Eberle

COLUMNS EDITOR
Anne Milley

TECHNICAL EDITOR
Daniel Valente

ASSOCIATE EDITOR
Amy Madison

ART DIRECTOR
Diana Witt

EDITORIAL DIRECTOR
Diana Levey

EDITORIAL AND TECHNICAL CONTRIBUTORS
Ian Cox, Landra Cunningham, Julia Gong, Atsuko Iono, Sachiko Ishibashi, Allison Lane, Geoff Mann, Arati Mejdal, Carla Paquet, John Sall, Maria Astals Serés, Audrey Shull, Taylor Sisk, Julie Stone, Isabelle Tamburro, Lou Valente, Jon Weisz, Trey Whittenton, Leo Wright, Rainbow Zhou

PHOTOGRAPHY
Steve Muir

WEB SUPPORT
Roger Alford

CIRCULATION
Amy Bullock, Antoine Smith

PRODUCTION
Melody Fountain, Larissa Muchnick

HOW TO REACH US
SAS Institute Inc. World Headquarters
JMP Division
SAS Campus Drive
Cary, NC 27513
919 677 8000
jmp.com

For US and Canada sales inquiries, call 877 594 6567.
JMP also has regional offices in China, France, Germany, Italy, Japan and the United Kingdom.

ABOUT JMP AND SAS

First launched in 1989, JMP is a software solution from SAS. John Sall, SAS co-founder and Executive Vice President, is the chief architect of JMP. SAS is the leader in business analytics software and services, and the largest independent vendor in the business intelligence market. Through innovative solutions, SAS helps customers at more than 80,000 sites improve performance and deliver value by making better decisions faster. Since 1976 SAS has been giving customers around the world THE POWER TO KNOW®.

This issue of *JMP Foreword* was published September 2016. The magazine also lives online: **jmp.com/foreword**

Luck is good, JMP® is better

I recently learned the word triskaidekaphilia. (Admit it, this one is new to you too.)

No, I didn't stumble upon it in my word of the day calendar. Rather, I was introduced to it by SAS co-founder and Executive Vice President John Sall, who also happens to be the creator of JMP.

With JMP now in its 13th version, the superstitious among us were concerned. In many places around the world, the number 13 represents bad luck and misfortune, but Sall believes that 13 is lucky for JMP. You'll learn why in this issue.

If you don't think you're inherently lucky, don't despair. The feature stories here demonstrate that luck isn't something that happens to you. It's something you make.

Lucky people are those who pay attention to patterns and maintain a healthy curiosity. And JMP is made for just this sort of person. JMP is for the inquisitive. For those who understand that real answers to real problems are there in the data, if only someone would use the right software to take a look.

You'll meet many people in this magazine who have made their own luck with JMP. Among them, Fresenius engineer Laura Zambianchi, who cites this Aristotle quote as a guiding principle: "Excellence is never an accident. It is always the result of high intention, sincere effort and intelligent execution. It represents the wise choice of many alternatives – choice, not chance, determines your destiny."

These pages are full of people who have determined their destiny with wise choices. Take the scientists at the San Diego Zoo who rely on findings from complex survival analysis to make strategic conservation decisions. Or the W. L. Gore engineers who transformed their experimentation process to improve their high-performance fabrics.

In fact, exploration geologist Steffen Brammer says his use of JMP to make decisions gives him "an unfair advantage." And who wouldn't want that? Forget a roll of the dice. Try JMP instead.

Jessica Marquardt
Editor

CONTENTS

COLUMNS

3 The luck of 13
John Sall, SAS co-founder and Executive Vice President, is confident that JMP 13 will improve your luck by giving you everything you need to make serendipitous discoveries.

5 My journey with JMP®: Peanuts to P&G leadership
What do peanuts, cookies and diapers have in common? **Tom Lange**, a retired R&D director from P&G, shares how he has used JMP to optimize products and influence decision makers.

7 How to negotiate experiment size when choosing a design
By exploring alternative design sizes, Los Alamos National Laboratory's **Christine Anderson-Cook** positions herself to ask for more resources.

10 Making sense of sensor data
With the internet of things, sensor data abounds. **Daniel Valente** and **Shannon Conners** on why JMP is the ideal sandbox for sifting through this data.

11 A health screening for your processes
The quality doctor is in! **José G. Ramírez** outlines how to give your processes a checkup.

13 Tomorrow is a brand new day: Text analytics in JMP® 13
Text data is typically unruly and difficult to analyze. But it can be done. Adsurgo's **Heath Rushing** introduces the new Text Explorer.

15 Protocol compliance oversight in clinical studies
Every drug developer has a regulatory – and ethical – requirement to manage the risks and benefits of clinical studies. Chimerix Medical Director **Enrikas Vainorius** on JMP Clinical for FDA compliance.

NEWS

18 Announcing JMP® 13

24 The power of Pro

FEATURES

28 It's survival analysis – literally
Scientists at the San Diego Zoo use analytics to give the burrowing owl its best chance to make a comeback.

32 Optimizing DOE for optimal product design
From tactical military shelters to waterproof jackets, W. L. Gore is the innovator behind the technology.

35 Changing the complexion of analytics
JMP is a clear choice for Nu Skin's promotional analysts, who have a wide range of language and technical skills.

37 Data analysis helps track patients in the HIV care continuum
An epidemiologist with the Colorado Department of Public Health finds those living with HIV/AIDS to avoid lapses in medical care.

39 The Beijing Aquarium finds clarity in analytics
Engineers optimize water quality conditions for a pod of beluga whales – and decrease energy use in the process.

42 From scraps to interface in six weeks
Medical device manufacturer Fresenius develops a tool for a complete, real-time picture of what's happening on the shop floor.

45 Pioneering scientific analysis
An exploration geologist uses cluster analysis to find high-nugget gold deposits.

47 Data-driven Quality by Design boosts drug development
Mitsubishi Tanabe Pharma seeks to offer state-of-the-art drugs that aren't cost-prohibitive.

49 Closing the skills gap
Cal Poly prepares graduates to keep pace in the age of big data.

52 Analytics cures a giant headache
Macfarlan Smith eases recurring processing pains and applies similar analyses to other projects.

55 Mining big data propels increased product miniaturization
Murata, a major supplier of advanced electronic components, champions data mining to achieve big process improvements.

57 Navigating road safety improvements
An industrial engineer at North Carolina State University studies ways to reduce car crashes caused by driver distraction.

59 Design of experiments brings order to the research cycle
Arlenda helps its pharmaceutical clients apply statistics from discovery to manufacturing.

The luck of 13

by John Sall

Thirteen is usually considered so unlucky that many buildings avoid having a 13th floor. But at JMP, we embrace 13 because we think that JMP 13 will improve your "luck." Here's why.

JMP's charter with serendipity

There's a word that combines luck with discovery: serendipity.

JMP has lived up to its tagline of being a great tool for discovery. From the beginning, we followed three principles:

First, JMP is very **graphical**, and the graphics empower discovery. With graphs you notice patterns in your data and points that don't fit patterns, both of which lead to discoveries. With graphs you have a better understanding of how well your data is supported by the models you try, and the graphs often suggest other models.

Second, JMP is very **interactive**, and interactivity allows you to follow clues, which lead to discoveries. Hover over a point to identify it. Use interactive commands to add more features to the analysis. Interactively edit your models. Exploration is enriched by interactivity.

Third, JMP is very **easy to use**. With an easy-to-use tool, you are encouraged to explore your data more, which, as with first two principles, leads to more discoveries. Ease of use keeps you in the flow, focusing on the revelations of your data, rather than navigating around the roadblocks of getting to the next step.

When combined, these three principles have made JMP a great discovery tool since the beginning of JMP.

When the power of experimentation is coupled with efficient experimental designs, you can discover how your processes behave – and how to improve them. The power of the analytical tools in JMP allows you to fit the data in many different directions.

So JMP helps discovery, which opens the door to serendipity and allows you to make your own luck.

JMP® 13 and serendipity

The new features in JMP 13 only increase your luck:

- JMP 13 brings data together more easily with **JMP Query Builder** and linked tables (virtual joins). Combining tables is important because data stored in relational databases typically normalizes the data into multiple tables to avoid duplicating any values and risking inconsistent data. JMP 12 made combining these tables easy, and now Query Builder is available for regular JMP tables, as well as tables from databases. With virtual join, you also have the option of keeping the tables separate while appearing as if they were joined, through the magic of links from one table to another through their keys.
- JMP 13 lets you analyze text data. **Text Explorer** parses columns containing text into words and analyzes them. Say you are looking for specific reference numbers that have syntax that can be described with a regular expression, Text Explorer makes it easy to find all those references.
- **Formula Editor** now allows you to find columns in a wide table much more easily.
- **Process Screening** simplifies searching through the reports of large-

John Sall, co-founder and Executive Vice President of SAS, joined Jim Goodnight and two others in 1976 to establish what is now the largest independent vendor in the business intelligence market. Sall designed and developed many of the software's earliest analytic procedures; he also created JMP software in the late 1980s and remains its chief architect. Sall is a Fellow of both the American Statistical Association and the American Association for the Advancement of Science.

Discovery and serendipity are all about being observant – and extending observations to good reasoning to hypotheses to analyses to conclusions. Having great tools to help you look at and model data makes everything so much better.

scale production data, allowing you to sort them by Stability or Capability so you can focus on the processes that most need attention.

- **Compare Designs** allows you to contemplate which designs are better at producing the effects you want in Design of Experiments.
- **MaxDiff** gives consumer research new power by providing individual estimates for choice experiments, supplemented with multidimension scaling.
- **Latent Class Analysis** adds new flexibility when clustering categorical variables or clustering variables instead of rows.

The serendipity story

Wikipedia offers this story about the origin of serendipity:

The King of Serendip had three sons. He exiled his three sons from the kingdom so that they would learn how to live without the shelter of royalty. The three princes were on a road far away when they noticed a few things on the roadside. What they saw led them to conclude that a camel had been there recently, one which was lame, blind in one eye, missing a tooth, carrying a pregnant woman, and bearing honey on one side and butter on the other.

Soon after, they told their suppositions to a merchant, who accused them of stealing the camel, although they had never even seen the camel. They were taken to the emperor of that land, where they explained the observations and deductions leading to their conclusions.

In this Persian Sherlock Holmes story, the three princes proved their power of observation, of using data to draw conclusions. The emperor absolved them and then appointed the princes as advisers.

The emperor was the ruler of the Sasanid Empire. (It's funny how SAS fits in the name, perhaps even serendipitous.) I will leave it to you to look up the princes' Sherlockian ratiocination about the camel.

Discovery and serendipity are all about being observant – and extending observations to good reasoning to hypotheses to analyses to conclusions. Having great tools to help you look at and model data makes everything so much better.

With JMP 13, we want to convert you from *triskaidekaphobia,* the fear of 13, to *triskaidekaphilia*, the love of 13. Good luck!

ONLINE

Watch Sall present the reasons why you will love JMP 13 at Discovery Summit:
jmp.com/discovery-sall

My journey with JMP®: Peanuts to P&G leadership

by Tom Lange

For me, it began in the late 1970s, way before JMP 1.0b. I was a brand new "out-of–school-kid" engineer who just started working for Procter & Gamble in Jif peanut butter. I was challenged to understand our peanut roasters. We asked seemingly simple questions like "how brown do the nuts get in a given time, at different air temperatures, with different amounts of air?" It would have made a great high school science project, except it was my job and the beginning of my career. And it turned out to be a serious engineering challenge. If we could understand it and optimize, it meant a lot of money for the company – not to mention making the best roasted peanuts possible for "choosy moms."

At first, things did not go well. Although the supporting statistician helped with well-designed first experiments (even by today's standards), when he did the regression it turned out that not one main effect among the main variables emerged as significant. The statistician said to me, "Look, there is a bunch of stuff going on here that is either grievous experimental error, or huge curvature, or I have no idea. I can't help you anymore. You are the engineer, you know this process, you have to dig around, do some different experiments and discover these behaviors yourself." On the stats side, he took the time to teach me some early new computer software on our mainframe called SAS® and SAS/GRAPH®.

From that day on, everything changed. I learned to do my own linear regression, then multivariate regression, and even nonlinear regression. I printed thousands of graphs with multicolor plotters, and then redesigned experiments. I had the time of my life! OK, maybe the time of my professional life. The story ended well for Jif and for me, and I never forgot the lessons of that experience.

> The JMP team works really hard to develop those killer visuals that tell the whole story in one glance.

After peanut butter, a bunch of us moved on to Duncan Hines chocolate chip cookies. The challenges were the same, they just tasted different. We were now doing regression on 1,000 test bakes with hundreds of cookies each, and ended up with a production-worthy simulation with 10 outputs and 40 inputs – things like "how much more flour in the dough does it take to make the diameter of a cookie smaller by .05 inches?" and "how does baking soda affect the color at the same temperature and bake time?" Since we did not have 30 years of experience as bakers, we had to accelerate the process with experiments and summarize them with models. In the end, the simulation helped us train the technicians and launch that product on time and with the highest quality.

OK, enough of really ancient history. Let's start catching up. Fast forward to 1989; I was in baby diapers then. No, not literally, but I was working on the Pampers brand.

I saw a beta copy of this software for a Mac (I was an early Mac adopter) called JMP 1.0b. That was so cool! It did not do everything, but what it did was amazing. Take, for example, how it handled nonlinear regression. It provided the partial differentials automatically, had a graphical input sheet and automatically

Tom Lange is a retired director of R&D at Procter & Gamble, where he founded and directed the modeling and simulation group. This team led efforts in consumer modeling, computational chemistry and biology, computer-aided engineering, and production system throughput and reliability. He now spends his professional time consulting with small and medium enterprises on ways to improve their competitive edge with the latest computer-based modeling and simulation tools.

These folks just want to know what time it is, not how to build their own clock.

spit out a comparison of the data and the model in an easy-to-understand graph. It was the same work we used to do, but it was so much easier and faster, as well as a lot easier to teach someone else.

That simple JMP 1.0b software reminded me of the early excitement I felt back in the peanut butter days, when I was learning so quickly. One morning, I was doodling around with a data set with a colleague, using JMP to plot how the amount of urine the child put in the diaper (urine load) affected the probability of a urine leak in an experimental diaper. I thought we should do an ANOVA and find out the average leaker "loading" versus the average non-leaker "loading."

Well, I got the x and y mixed up accidentally and put the Urine Load in the X (continuous) and the Leak/No Leak in the Y (nominal), and out came all this stuff (no pun intended): an S-shaped curve, a formula predicting the probability of a leak as a function of the urine load, a p-value. It turned out to be logistic regression, before PROC LOGISTIC even was available in SAS 6. I used the question mark tool from the menu, and it taught me quite a bit. I went on to study that approach in some detail later in the literature, but I would not have even known what to look for without JMP 1.0.

As JMP 1 progressed to JMP 2 and JMP 3, I made yearly visits to North Carolina and John Sall. P&G would argue for including things from simple calculator features to full-blown reliability modules. In later JMP versions, our logistic experience led us to Kaplan-Meier reliability analysis and ultimately a lot of the reliability analyses available today. The relationship between P&G and the JMP team is still growing, and I am proud to say my last day with P&G was spent in Cary, NC, visiting SAS with my P&G replacement.

In those last 25 years, JMP grew to be much more than a personal analysis tool. I went to the "dark side" and became a manager, eventually becoming Director of Modeling and Simulation in 2004. I found I was able to influence the organization through hiring, promoting modeling and simulation, and fostering a culture that supported analytics.

As leaders, we came to expect every engineer and scientist to do their own data exploration, their own experimental design and, most importantly, their own storytelling of their results.

JMP as a storytelling aid may be its most important contribution. The JMP team works really hard to develop those killer visuals that tell the whole story in one glance.

With JMP, you can lead your audience to that final visual in real time from the raw data and analysis. Don't just present the final answer and then spend 10 minutes trying to explain where it came from – show them where it comes from. Let them have a taste of that discovery experience! For example, you can start with a table of measurements (a big one, scroll around top to bottom, etc.), plot a simple histogram, then fit that with a normal or lognormal or whatever curve, test for its normality, put in specification limits and calculate Cpk, out of limits and do that live in a presentation. If someone in the audience has a question you had not thought of, chances are you can probably accommodate the calculation right on the spot and – most importantly – give a graph to tell the story.

Influencing nontechnical decision makers may be one of the most important skills we as scientists and engineers must master, whether it's telling folks when to evacuate, showing people something is safe, or telling management what will likely happen. These folks just want to know what time it is, not how to build their own clock. The graphics, the ability to tell the story in real time and with a memorable visual, continue to be one of the most important features of JMP.

Here we are, at the introduction of JMP 13, and the buzz of what's new in this one is creeping down my back again – especially for categorical data, something that has been so hard to analyze correctly for years. But it is so important for shaping decisions when numerical metrics are hard to come by. So here is my challenge for you: Unwrap JMP 13 and find out something new in your data. And have the time of your life (OK, career life) doing it!

ONLINE

Tom Lange talks about building an organizational culture of analytics on Analytically Speaking: **jmp.com/lange**

Improve your process with statistical models: **jmp.com/improve**

See how you can convey results with data visualization: **jmp.com/dataviz**

How to negotiate experiment size when choosing a design

by Christine Anderson-Cook

In classes where we learn about design of experiments, the problem is often posed as "Select a design with N runs …." So it is tempting to think about this as a given constraint in our selection process. However, it can be to our great advantage to think about the size of design as a flexible criterion on which we want to evaluate our choices and make comparisons. In recent years, I have been an active advocate for looking at multiple criteria when examining the appropriateness and desirability of designs for the goals of our experiment.[1,2,3]

Classical designs like factorial and fractional factorial designs for first-order models, and central composite or Box-Behnken designs for second-order models, have long had an important role in design of experiments if we are interested in spherical or cuboidal design spaces and if the size of our design matches the standard sizes. More recently, computer-generated D- or I-optimal designs have become available to allow greater flexibility when the shape of the design space and/or design size are nonstandard. Both of these categories of designs have strengths and weaknesses, so using the Evaluate Design platform in JMP allows careful exploration of important questions such as:

- What kind of power can I expect from the design?
- Will I be able to estimate all the terms in my model independently?
- If the model is misspecified, how will parameter estimates be affected?
- How well will I be able to predict the response in different regions of the design space?

Understanding these important questions before the experiment is run can save us from running an ineffective experiment that cannot do what we want it to, or allow us to choose between alternatives to match the priorities of our experiment.

For many years now, when I have been asked to design an experiment as part of a project, I have created multiple designs that not only included the target design size, but also a couple of choices that are smaller and larger. Since most studies involve questions that cannot be answered with a single experiment, it is valuable to view the management of resources in the context of sequential experimentation. If we can save some resources early in the process, that frees them up for potential later experiments. If we need more resources early to really understand what is going on with our process or product, this can position us to make key decisions and understand the underlying relationship between inputs and outputs. We can then use this knowledge to guide later experiments. Hence, the exploration of multiple design sizes to understand our choices.

In JMP 13, the Compare Designs platform makes these between-design-size evaluations easier. Not only can it make comparisons between different design choices of the same size, but it can also provide useful information about different-sized designs. Imagine that you are asked to create an 18-run design to estimate the main effects for seven factors. Constructing the right designs to compare is also easy with the Custom Design and Augment Design platforms.

It can be to our great advantage to think about the size of design as a flexible criterion on which we want to evaluate our choices and make comparisons.

Christine Anderson-Cook, PhD, is a Research Scientist at Los Alamos National Laboratory, leading projects on complex system reliability, non-proliferation, malware detection and statistical process control. She has authored more than 200 articles in statistics and quality journals, and has been a longtime contributor to the Quality Progress Statistics Roundtable column. Anderson-Cook co-authored a popular book on response surface methodology and is an elected Fellow of the American Statistical Association and the American Society for Quality. She is the recipient of the 2012 William G. Hunter Award, and a two-time recipient of the ASQ Shewell Award.

You could use the Custom Design platform in JMP to construct designs with 16, 18 and 20 runs, and perhaps also include a definitive screening design (with a default of 17 runs). Armed with these four designs, we can now consider the impact of design size on their performance. Note that JMP 13 allows comparison of up to three designs in a single window, but it is easy to have two Compare Design windows open with the same base design to allow for easy comparisons of up to five designs. Table 1 describes some questions of interest with the JMP tools that make it easy to assess these aspects of the designs.

Category	Question	JMP Summary
Power	Depending on how large we anticipate the main effects to be, do we need the additional power of the larger designs?	Power Analysis Power Plots Power vs. Sample Size
Ease of estimation of model terms	Are there important differences in the correlation structure of the terms in the model?	Color Map on Correlations Absolute Correlations
Aliasing	Are there differences in the correlation structure between terms in the model and potentially active two-way interactions?	Alias Matrix Color Map on Correlations Absolute Correlations
Prediction variance	What are the overall differences in the prediction variance change between the different designs? Are there regions with different prediction variance performance in the design space?	Fraction of Design Space Plot Prediction Variance Profiler and Surface

Table 1: Aspects to compare between competing designs

Figures 1 through 3 show a sample of the tools available in JMP 13 for the three (16-, 18- and 20-run) custom designs with the 18-run chosen as the base design. In Figure 1, I have adjusted the different factors to have different sized anticipated coefficients to see the ranges of powers. (Note: These designs are all symmetric with the same power for all effects of the same size.) Figure 2 provides one summary of the correlations. You'll see there are interesting comparisons to consider: The 18-run design has some slightly correlated main effects in the primary model, while the 16-run and 20-run designs have no correlation between any terms in the primary or possible models. It is not uncommon to have some sample sizes that are less suited to good balance and symmetry. The 16-run design has the smallest correlations between all sets of terms considered in Figure 2, but it has slightly less power than the 18- and 20-run designs. Figure 3 shows a summary of the overall prediction variance profile across the seven-factor design

Power Analysis

Significance Level 0.05
Anticipated RMSE 1

Term	Anticipated Coefficient	18-run design Power	16-run design Power	20-run design Power
Intercept	1	0.963	0.937	0.983
X1	0.5	0.471	0.421	0.538
X2	0.75	0.806	0.748	0.868
X3	1	0.963	0.937	0.983
X4	1.5	1	0.999	1
X5	2	1	1	1
X6	3	1	1	1
X7	4	1	1	1

Apply Changes to Anticipated Coefficients

Good 0.8 0.6 0.4 0.2 Bad

Figure 1: Comparison of which effects are likely to be detected from 16-, 18- and 20-run custom designs

Absolute Correlations

Model x Model	Average Correlation	Pairwise Confoundings	Pairwise Terms
18-run design	0.046	0	21
16-run design	0	0	21
20-run design	0	0	21
Model x Alias	**Average Correlation**	**Pairwise Confoundings**	**Pairwise Terms**
18-run design	0.152	0	147
16-run design	0.102	0	147
20-run design	0.167	0	147
Alias x Alias	**Average Correlation**	**Pairwise Confoundings**	**Pairwise Terms**
18-run design	0.11	0	210
16-run design	0.071	0	210
20-run design	0.111	0	210

Good 0.1 0.3 0.5 0.7 Bad

Figure 2: Comparisons of correlation structures between designs

space.[4] In this case, there seems to be a similar reduction in the prediction variance with each increase in sample size. How we value the differences between the designs is dependent on our priorities and our tolerance for different costs.

Once we have examined these characteristics of competing designs, we can make the right choice for our particular study. In addition, it will be much easier to defend our choice of design, because we understand what performance is realistic to expect before any resources are spent. If we need to make the case that a larger experiment is needed, this discussion will be concrete and driven by available information. If we are able to suggest a smaller experiment than originally requested, then we may buy leverage for the next experimental problem when perhaps more resources are needed early in the process. Whenever we can translate a fix constraint into a tactical choice made for a justifiable reason, we are going to feel much more comfortable about our choice.

Many times, the initial resources allocated to the experiment are not sufficient to realistically meet our goals. By exploring alternative design sizes, we can put ourselves into a better negotiation position for getting additional experimental resources. How would you rather state your case: "A bigger experiment would be better," or "The current experiment is not adequate for the power and prediction variance that we need to be able to answer the questions that the study is tasked to tackle. Here is an experiment that satisfies our needs"? We might not always win that negotiation and be able to increase the size of the design to our desired target, but we are making a compelling, data-driven argument that highlights the consequences and trade-offs of several alternatives.

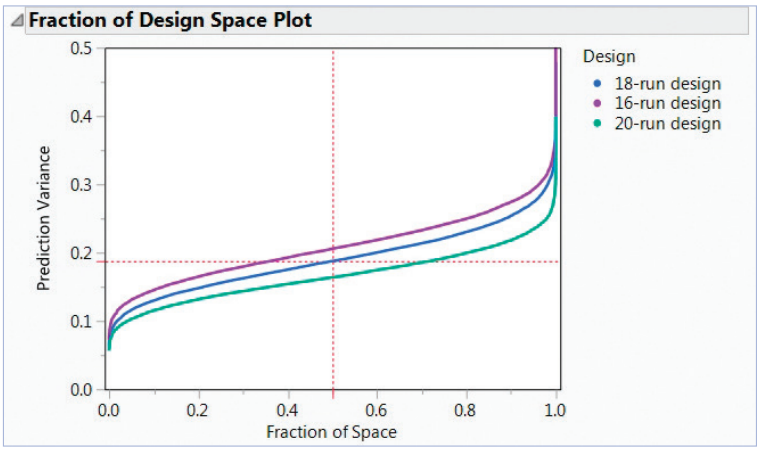


Figure 3: Comparison of prediction variances between designs

References

1. Lu, L., Anderson-Cook, C.M., Robinson, T.J. (2011) "Optimization of Designed Experiments Based on Multiple Criteria Utilizing a Pareto Frontier," **Technometrics** 53 353-365.
2. Lu L., Anderson-Cook, C.M. (2012) "Rethinking the Optimal Response Surface Design for a First-Order Model with Two-Factor Interactions, when Protecting against Curvature," **Quality Engineering** 24 404-422.
3. Lu., L., Anderson-Cook, C.M., Lin, D. (2014) "Optimal Designed Experiments Using a Pareto Front Search for Focused Preference of Multiple Objectives," **Computational Statistics and Data Analysis** 71 1178-1192.
4. Zahran, A., Anderson-Cook, C.M. and Myers, R.H. (2003) "Fraction of Design Space to Assess the Prediction Capability of Response Surface Designs," **Journal of Quality Technology** 35 377-386.

ONLINE

Christine Anderson-Cook shares research on objectively considering trade-offs: **jmp.com/anderson-cook**

Want to optimize your processes with DOE? Get the white paper: **jmp.com/optimize**

Compare designs with JMP: **jmp.com/compare**

Making sense of sensor data

by Daniel Valente and Shannon Conners

JMP can solve most data problems faced by scientists and engineers working with high-density sensors. From industrial devices living on the internet of things to health and fitness monitors, JMP is the ideal sandbox for sifting through noisy data to find important patterns.

Database Query Builders. For sensor data in structured SAS or ODBC-accessed databases, quickly create queries that join multiple tables, apply filters, and preview the results. Queries can be saved, shared and updated directly from an imported table. Move faster from raw data to insight by adding summary columns and graphics scripts to run automatically post-query.

JMP Query Builder. Even if you don't have direct access to your database, you can still use the query, filtering, summarization and sorting tools in Query Builder by importing a set of flat files into JMP tables. And save the query as a script to rerun automatically. Just pull the most recent data, rejoin, and move forward with analysis.

New Formula Column. Often, sensor data is collected and stored with a time stamp, sensor name and sensor value to conserve storage space. But pairing raw data with summarized values helps you answer broader questions, clean up artifacts and discover trends. Right-click any data column to add a new formula column. The formula shortcut menu offers quick access to the Group By functionality that identifies a column of categories for calculating new measures.

Virtual Join. When working with sensor measurements, you may not have the required memory to merge summary values directly into a raw data table. Use virtual join to link values stored in separate tables for better analysis and graphing. Link and unlink data sources on demand, making it easier to update disparate data sources independently.

Process Screening. The Process Screening platform does the upfront heavy lifting of consolidating control chart statistics and recent shifts into a single table that you can sort to find important process variables, allowing you to focus your efforts and drill down to critical control charts.

Modeling and Cross-Validation. JMP includes a full suite of linear and nonlinear modeling tools that build useful models on their own but also link into cross-validation tools, which is crucial for evaluating observational data.

Graph Builder. When exploring sensor data, you need to quickly create and customize a variety of multi-element graphs to spot patterns and potential artifacts. JMP Graph Builder provides many graph types and an intuitive drag-and-drop interface so you can save and reproduce graphs easily.

Selection Filters. JMP provides traditional list-based filters for exploring custom data slices; you can also use a summary graph as a filter for one or more graphs. Simply select elements in your filter graph to refocus your report on the area of interest.

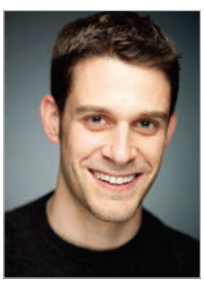
Dashboards. By showing patterns in your data in a concise, consistent report, dashboards help you communicate key findings to decision makers. Create a new dashboard from a template, drag and drop tables and reports onto the canvas, save and refresh your dashboard as needed, or export it as an interactive HTML web report for colleagues without JMP.

Sharing with Scripts and Add-Ins. When you construct a workflow that gives you the results you need, you can easily share your detailed steps with other JMP users. JMP autogenerates scripts to reproduce data manipulation steps and custom graphs. You can also develop custom menus and add-ins to walk colleagues through complex workflows.

Production systems are great at telling you what is going on right now – helping you run business rules, prevent dangerous conditions and operate systems within predefined spec limits. But a world of insight also lives in your historical data. JMP is a no-risk sandbox for finding patterns and trends in data.

ONLINE

Read the full article on the JMP Blog:
jmp.com/sensors



Daniel Valente, PhD, is a JMP Product Manager and **Shannon Conners, PhD,** is a JMP Development Director at SAS. They use JMP outside of work too, analyzing quantified home and quantified self data.

A health screening for your processes

by José G. Ramírez

We strive to achieve processes that are lean, capable and stable, but how do we know how healthy our processes are? How do we know if they are meeting customer requirements? How do we know if they are stable or predictable?

Process performance metrics

The centered capability index, Cpk, and the centered performance index, Ppk, are popular metrics that tell us how much elbow room is available within the space defined by the specifications. In other words, they tell us how much our process can move before we produce nonconforming output. What has been missing, however, is a way to also measure how stable, or predictable, our processes are. Yes, it is important to know if our processes are capable, but this is just a one-dimensional view of our processes. A capable process that has an unpredictable output can become incapable.

Like fractals, a predictable process is self-similar in the sense that both a global and a local measure of its output variation should give similar results. If that is the case, then the ratio of the global variation to the local variation should be close to one. This is exactly what the Stability Ratio (SR), introduced by B. Ramírez and Runger in 2006, does. It compares a global measure of variation – the overall variance – with a local measure of variation, the variance derived from a process behavior chart. Now we have a metric to evaluate how predictable a process is: If the process is predictable, or self-similar, the overall and local variances are similar, and the SR is close to one.

Process screening

Just like health screenings are performed to identify diseases or possible medical conditions, a process screening is a way for us to understand how well our processes are performing. The good news is that in JMP 13, there is an app – I mean, a platform – for that: the Process Screening platform. The new Process Screening platform helps us perform a fitness evaluation of multiple processes all at once, which is especially helpful when you are dealing with hundreds – or even thousands – of process variables and you need to quickly identify the problematic ones. It also provides many useful metrics to assess the health of your processes, including the SR metric to evaluate stability. To my knowledge, JMP 13 is the only statistical software that currently offers the SR metric.

Let's say you have seven processes, A through G, that you want to evaluate, or put through a "health screening." We want to know if your processes are capable and stable (predictable), and if there are any areas for improvement. Here is the first half of the Process Screening platform output (Figure 1).

The first column is the process name, while the second column shows the overall alarm rate for each process, or

José G. Ramírez, PhD, is a statistical engineer with years of experience in the semiconductor, electronics and biotech industries. A JMP user for more than 25 years, he works closely with engineers and scientists to help them make sense of data, and through collaborative education, helps promote statistical thinking and JMP usage.

Process Screening

			Western Electric - Nelson Rules								Range Limit	Latest
Column	Alarm Rate	Any Alarm	Test1	Test2	Test3	Test4	Test5	Test6	Test7	Test8	Exceeded	Alarm
E	0.20625	33	4	20	0	0	4	9	0	0	5	25
C	0.08750	14	4	6	0	0	4	2	0	0	4	28
G	0.08750	14	2	0	0	0	0	1	8	0	4	32
F	0.06875	11	3	0	0	0	1	1	4	0	4	15
A	0.06250	10	2	2	0	1	0	4	0	0	3	3
D	0.06250	10	3	0	0	0	1	0	4	0	5	3
B	0.03125	5	2	0	0	0	1	0	0	0	4	23

Figure 1

what Ramírez and Runger call the instability ratio (I_{NSR}). In this report, the alarm rate is composite of all the failure rates from the eight Western Electric tests that are applied to the process behavior chart, but the number of tests applied can be customized. Right away, we can see that Process E has an overall alarm rate of 20 percent, with most of the alarms (20 of them) coming from Test 2. The output also shows that the latest alarm comes from subgroup 25, which is helpful information for troubleshooting.

The other half of the output is where we find the capability and stability pulse of our processes. For example, for Process E the SR=1.45 and the Ppk=1.93, while for Process B the SR=0.97 and Ppk=1.85. The high SR for Process E indicates a stability issue, since for a stable process the SR~1, but the large Ppk indicates that the process is capable (Figure 2).

Process Screening

Column	Within Sigma	Overall Sigma	Stability Ratio	Mean	Count	Subgroups	Capability: Out of Spec Count	Capability: Out of Spec Rate	Capability: Latest Out of Spec	Capability: Cpk	Capability: Ppk
E	0.05078	0.06116	1.45	1.14506	160	160	0	0	.	2.330	1.934
C	4.26907	5.42938	1.62	41.0884	160	160	3	0.0188	29	1.086	0.854
G	0.06109	0.06514	1.14	1.22544	160	160	0	0	.	1.776	1.665
F	0.06037	0.06934	1.32	1.21954	157	160	1	0.0064	105	1.549	1.348
A	3.93996	4.15997	1.11	76.065	160	160	0	0	.	1.602	1.517
D	4.61691	5.51924	1.43	45.4693	160	160	0	0	.	2.561	2.142
B	5.37716	5.29746	0.97	85.5725	160	160	1	0.0063	24	1.824	1.852

Figure 2

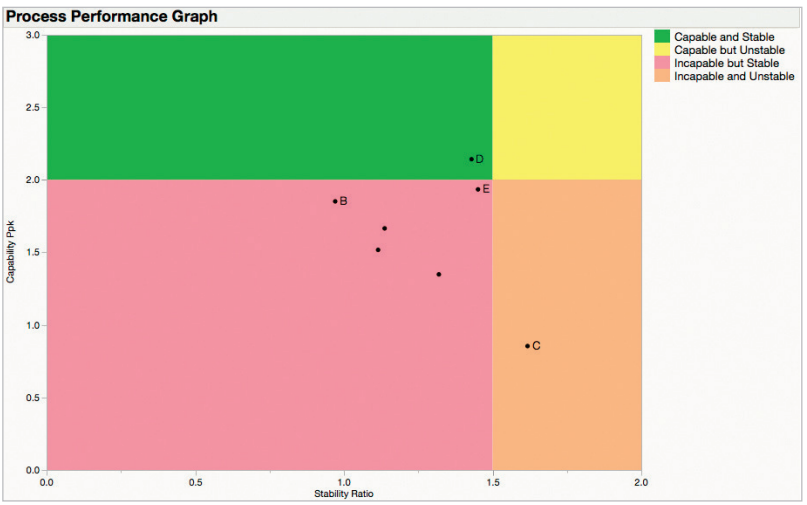


Figure 3

Visualizing process performance

From the Process Screening output we can quickly get a sense for each process, but what if we want to visualize all this information at once? The Process Performance Graph within the platform displays the SR on the x-axis and the Ppk on the y-axis. The four quadrants in the graph help us understand how our processes are performing in the two dimensions of stability and capability. Processes in the green area are capable and predictable; this is the ideal or no-problem state. Processes that land in the red area are processes with yield issues, that is, predictable but not capable; and processes in the yellow area are capable but unpredictable. Finally, processes in the orange area are in "double trouble," that is, unpredictable and not capable. Wheeler (2010) has a good discussion on these four stages, how entropy pushes processes to the double-trouble state, and what to do to achieve the ideal state (Figure 3).

By default the Process Performance Graph has boundaries at Ppk=2, corresponding to a 6σ quality level, and SR=1.5. A quick glance at the graph shows that five out of our seven processes are in the red area, corresponding to yield issues, including processes B and E, which we discussed above. One process, Process C, is in the double-trouble area, while Process D is in the ideal state.

Knowledge and efficiency

The Process Performance Graph gives us a process-health snapshot for a given time period, but things change over time and, like health screenings in general, we need to assess our processes on a regular basis. Comparing these process-health snapshots over time will reveal how our improvement efforts are working against the natural system entropy, and how, due to these efforts, our processes move from the orange, red and yellow areas to the green area, or the ideal state of capability and predictability.

The Process Screening platform is an efficient way to perform a health screening on multiple processes at once, and it will greatly reduce the time required to perform these analyses.

ONLINE

See José Ramírez discuss quality engineering with Brenda Ramírez:
jmp.com/ramirez

Read his book *Analyzing and Interpreting Continuous Data Using JMP*:
jmp.com/continuous

Use JMP for quality engineering:
jmp.com/quality

Tomorrow is a brand new day: Text analytics in JMP® 13

by Heath Rushing

December 2001. I am in a room with two colleagues, analyzing text from thousands of nonconformance reports. After reading each root cause, we categorize each nonconformance into "buckets" of systemic problems across multiple sites. I was a quality engineer for a large biotech company that, like many FDA-regulated companies, wanted to ensure that analysis of nonconformances was the foundation for their corrective and preventive action (CAPA) program. The high-frequency/impact categories would receive top priority for improvement projects in 2002. We worked on this for three weeks. Although this was a critical project for the company, our method of reading each nonconformance, then classifying it into a category, was highly inefficient.

March 2016. I am in the office of one of those former colleagues who is now working for a large medical device company. We begin to talk about the enormous amount of unstructured (text) data that his company collects daily: nonconformances, product complaints, maintenance reports, customer feedback reports, even employee satisfaction surveys. He laments that the analysis of this unstructured data is so cumbersome that analysis of text data not tied to compliance is largely ignored. I reply, "If you are not going to analyze it, then why collect in the first place?" If only there were an efficient way to quickly analyze text from these documents to highlight the top priorities from each of the different functional areas (regulatory, compliance, maintenance, marketing and HR), allowing future decisions to be more data-driven.

Now there is a way, with the Text Explorer platform in JMP 13.

I used a familiar data set on FDA inspection citations from medical devices companies to illustrate the utility of text analytics in JMP 13 to my colleague.

From 2006 to 2012, there were 27,594 inspection citations, each containing text describing the audit observation. We used this collection of documents (often called a corpus) to find the most frequently cited audit observation themes from medical device companies.

Within seconds, he identified five prevailing themes from the Terms and Phrases Lists: procedures have not been adequately established/implemented, CAPA, inadequately documented, design controls (inputs, outputs, verification, review, history file), and complaint handling procedures.

Because a picture is worth a thousand words, I used a word cloud to display the most frequent terms and phrases in the corpus (Figure 1). Terms get smaller as they become less frequent. The orientation of terms is irrelevant.

As you can see, words from his initial list of the five most frequently cited

Heath Rushing is the co-founder of Adsurgo and author of the book *Design and Analysis of Experiments by Douglas Montgomery: A Supplement for Using JMP*. Previously, he was the JMP and Six Sigma training manager at SAS. As a Quality Engineer at Amgen, he designed and delivered a DOE course that immediately became the company standard required at multiple sites. He won the top teaching award out of 54 instructors in the US Air Force Academy math department, where he taught several semesters and sections of OR and statistics. Rushing has been an invited speaker on applicability of statistics for national and international conferences. Additionally, he designs and delivers short courses in statistics, data mining and simulation modeling for SAS.

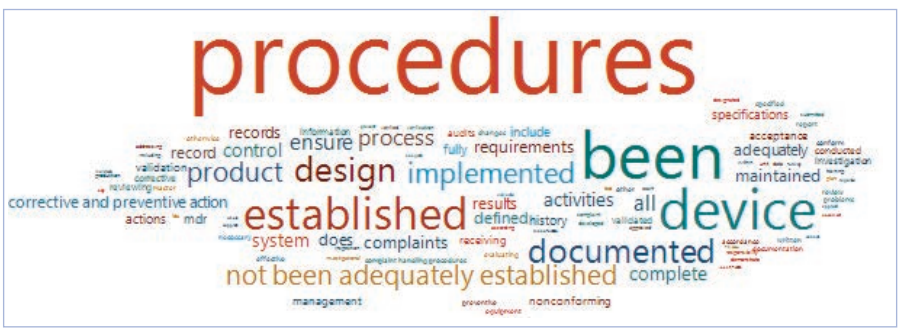


Figure 1: Word Cloud

observation themes appear prominently in the word cloud.

Although the Terms and Phrases Lists and word cloud provided him the ability to quickly construct an initial list of frequently cited observation themes, he could have also automatically assembled frequently used words and phrases. Text Explorer allows you to find additional information from the corpus by clustering words: determining what sets (or clusters) of words are used together. We constructed 100 clusters of words from the corpus. From this sorted list, we found that *appropriate*, *statistical* and *technique* clustered together, as well as *unresolved* and *discrepancy*. This is indicative that companies were cited for not documenting appropriate statistical techniques used and did not adequately resolve discrepancies. Although word clustering is useful, it does not provide an indication of the observed frequency for these clusters of words.

JMP 13 Text Explorer first organizes all the words as columns and all the documents as rows of a large matrix called a Document Term Matrix (DTM). The matrix elements are the number of times (frequency) each word occurs in the particular document – usually 0, hence a sparse matrix. It then develops numerical representations for the relationships between words, as well as documents. This is called Singular Value Decomposition (SVD). By plotting different singular value vectors versus others, you can graphically see relationships between words and documents, as well as an indication of observed frequency.

Text Explorer allows users to change the method each word is "weighted" in the DTM rather than just the raw frequency of word counts per document. This feature allows you to see relationships between words that may not have been evident in the Terms and Phrases Lists.

As can be seen from the Term SVD Plot (Figure 2), *incomplete risk analysis* and *documentation* on *equipment*

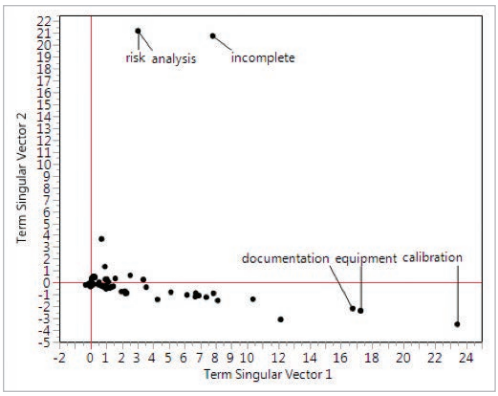


Figure 2: Term SVD Plot

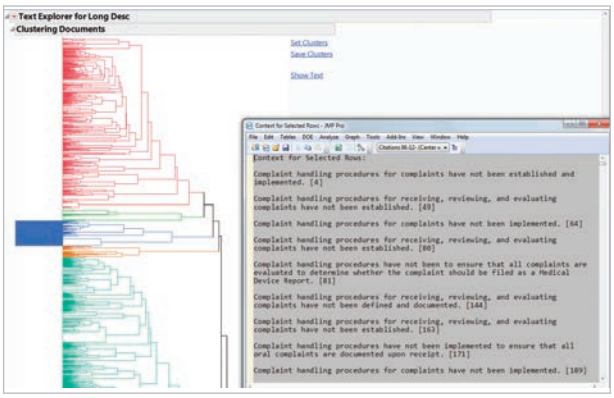


Figure 3: Clustering Documents Report

calibration are other frequently cited observation themes. You can further explore the most relevant themes in the corpus using topic extraction.

Lastly, we turned our attention to our 15-year-old dilemma – how to quickly categorize many different documents to find systemic problems across multiple sites. This is called document clustering. To illustrate the utility, I showed him how to cluster inspection citations. Clusters of documents will contain similar words and phrases used together.

The Clustering Documents report (Figure 3) shows that documents clustered together had violated statutes in the same (or similar) section(s) of the code of federal regulations applicable to medical devices (21 CFR 820). By selecting Show Text, we could immediately pull up the full descriptions of the text we were exploring; this is true not only in the Clustering Documents report but in any table or graph including the word cloud. Since the descriptions of the statute violation are purposely written with either the same or similar phrases, it came as no surprise these documents were clustered together. However, this same method can be used to cluster documents without the same wording.

Very quickly, my colleague learned the utility and efficiency of text analytics. In the past, JMP users have efficiently analyzed continuous and categorical data, even small amounts of text. However, analysis of large amounts of unstructured text data was unattainable. Tomorrow is a brand new day. Text Explorer is here in JMP 13.

ONLINE

See Text Explorer in action: **jmp.com/text**

Heath Rushing discusses text exploration on an episode of Analytically Speaking: **jmp.com/rushing**

Protocol compliance oversight in clinical studies

by Enrikas Vainorius

Under US Food and Drug Administration regulations, sponsors are responsible for ensuring the proper oversight of clinical studies. This oversight is a multifaceted and complex process, spanning from the investigator selection to contract organization management to overall data quality.

As a medical monitor, JMP Clinical helps me address one small portion of this elaborate process – compliance with the study protocol.

I would like to share my experience on a few topics for which I use JMP Clinical to simplify and expedite the review tasks, avoiding costly (in time and money) custom programming. This includes:

- Oversight of properly enrolled subjects (e.g., inclusion/exclusion criteria).
- Oversight of the safety of enrolled subjects (e.g., prohibited medications, study drug stopping criteria).

Do we have the proper subjects enrolled?

The success of an ongoing study starts with enrolling the appropriate subjects, as defined in the protocol inclusion/exclusion criteria. There is a reason why, in every study, we select a particular subject population – hence the name "controlled" clinical study. We rely on the appropriate training of the investigators and their staff to ensure that only the subjects who should be enrolled are. However, appropriate training does not dismiss the sponsor from the future oversight and assessment of additional training needs. In many cases, enrolling unqualified subjects not only jeopardizes the

Enrikas Vainorius, MD, is Medical Director of Clinical Research at Chimerix. He has more than 15 years of biomedical and clinical research experience obtained while working in different organizations, including hospitals, the pharmaceutical industry and academia. His experience in biomedical research ranges from disease physiology to the development of novel cellular therapies and diagnostic methods based on adult stem cells. During his work in the pharmaceutical industry, Vainorius acquired expertise in clinical trial monitoring and management, including clinical trial design and execution, medical and data quality oversight, data analysis, regulatory/procedural compliance with the FDA and other regulatory entities.

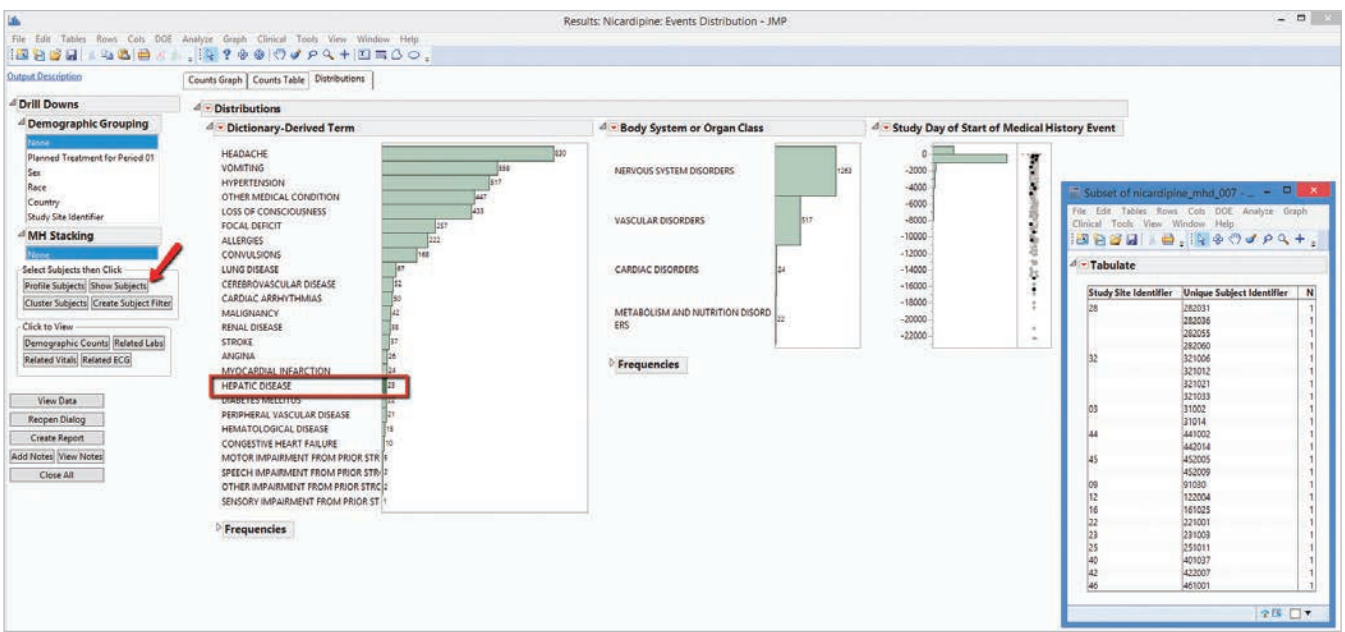


Figure 1: Drill down to see which sites have enrolled subjects with exclusionary pre-existing conditions

subjects' safety, but also has an impact on the analyses as defined by "per protocol population."

Of course, the frequency of reviews (e.g., how often data is delivered from data management vendors) and what action is taken (e.g., when retraining is warranted) will depend on the sponsor's internal procedures.

For demonstration purposes, let's say our protocol should exclude subjects with pre-existing hepatic disease. Using JMP Clinical, not only can I easily determine if any such subjects were enrolled, I can also identify exactly who those subjects are, and at which sites. As you can see in Figure 1, by drilling down I can see the data I need to have an informed discussion with at least a few sites that appear to have deviated from those exclusion requirements on multiple occasions.

Figure 2 presents another example of how we simplify the identification of issues and enhance our ability to react earlier. Of course, this will all depend on your data availability and review cycle.

Is the safety of subjects maintained during the trial?

Every drug developer has a regulatory – and ethical – requirement to manage the risks and benefits of the drug under development. Depending on the therapeutic area, subject population or other factors, clinical studies may have well-defined safety monitoring criteria, including the study drug-stopping rules defined in the protocol.

JMP Clinical helps to monitor such instances to confirm whether appropriate actions were taken for those subjects. If not, it assists the sponsor in taking appropriate action in a timely manner (e.g., re-training of the site[s], or closing them).

It is very common for certain medications to have an interaction with the study drug, which might increase or decrease the exposure of the study drug to the subject. Thus, in many protocols, we will find a prohibited medication section. Also, it is not uncommon that certain medications are allowed prior to the first dose but prohibited as a concomitant medication.

Figure 3 demonstrates how we can distinguish which instances of certain medications were given prior to the first dose and which ones (if any) were given (and how many times) in violation of protocol. In this example, oxycodone was not prohibited in the sample JMP Clinical database, but selected merely to demonstrate the process of identifying medication that was given after the first study drug dose.

We can also review the subjects to see if they indeed interrupted the study drug

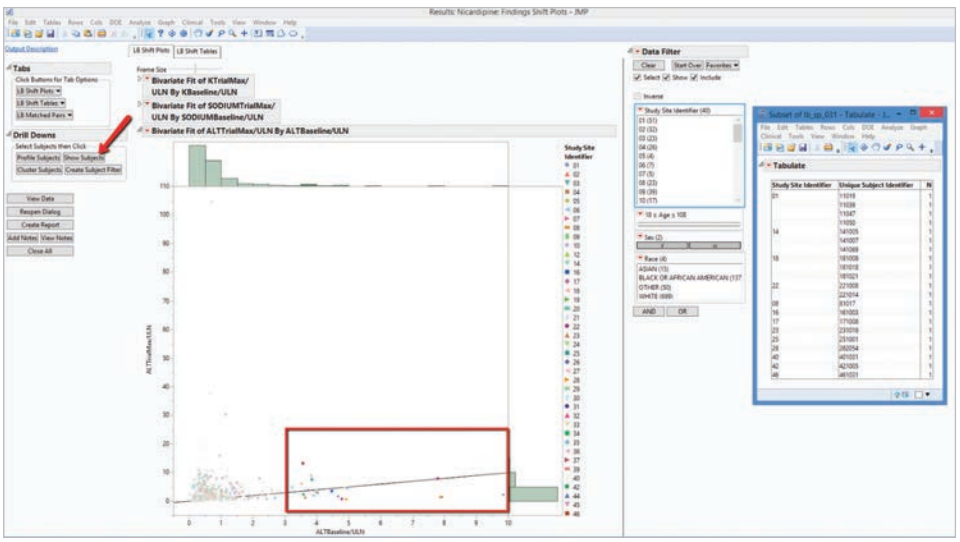


Figure 2: Identify the subjects who are enrolled outside the protocol requirement (to be excluded if baseline ALT is greater than 3 times ULN), along with their sites

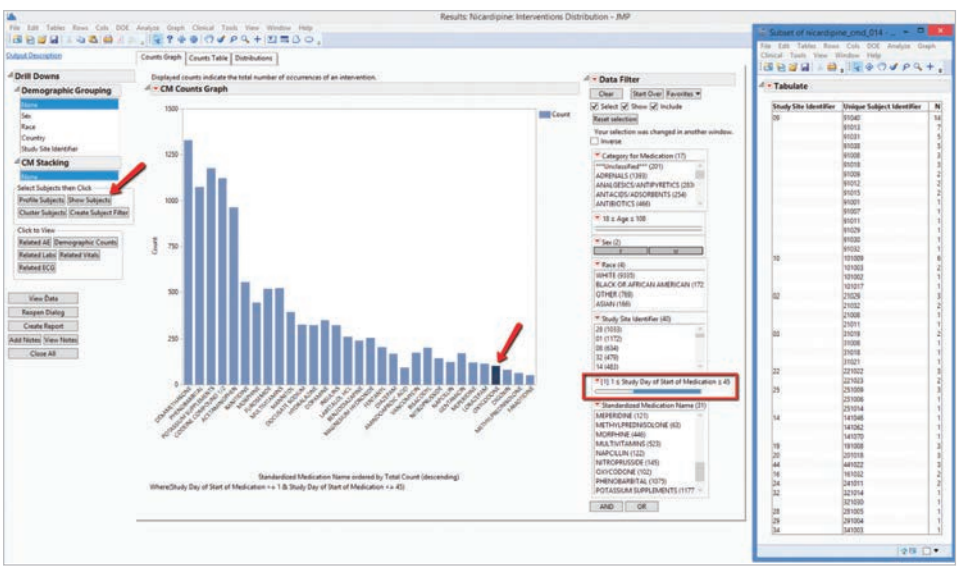


Figure 3: Find subjects who were given prohibited medications during the trial that might interact with the study drug

based on the ALT values (if not, it might require action). After identification of such subjects (Figure 4), we would need to decide if proper action was taken with the subject. One way to do this is to create a subject profile and review the parameters of interest. In this case (Figure 5), we would conclude that appropriate action with the study drug was taken for the highlighted subject.

These are just a few examples of how JMP Clinical aids clinical trial oversight for medical monitors.

ONLINE

JMP Clinical simplifies data discovery, analysis and reporting in clinical trials. Learn more: **jmp.com/clinical**

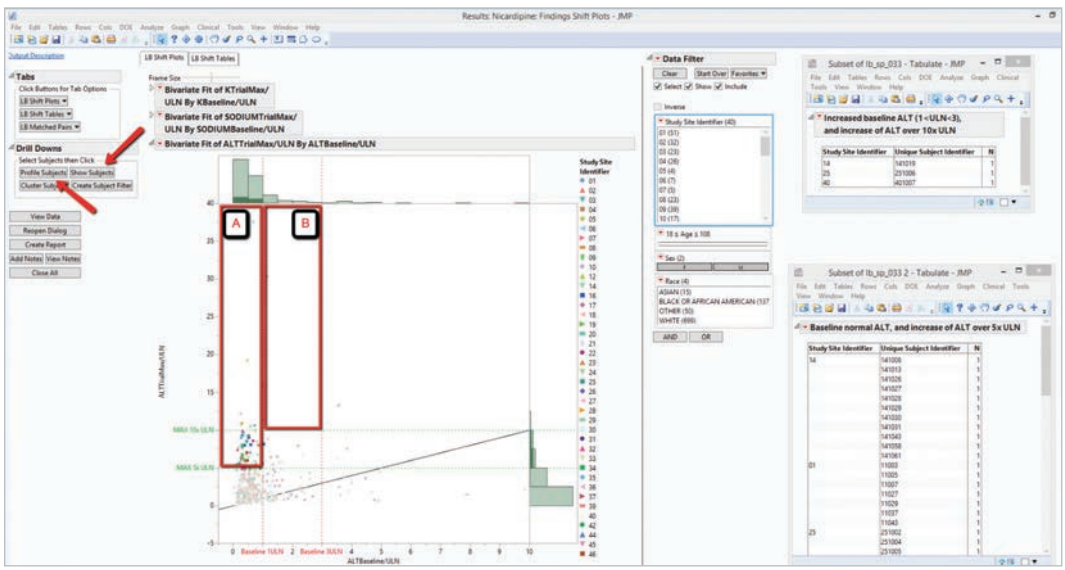


Figure 4: In this example, the study drug interruption should occur for subjects who had a baseline normal ALT and an increase of ALT over five times ULN during the trial (A), and for the subjects who had increased baseline ALT (1<ULN<3) and an increase of ALT over 10 times ULN (B)

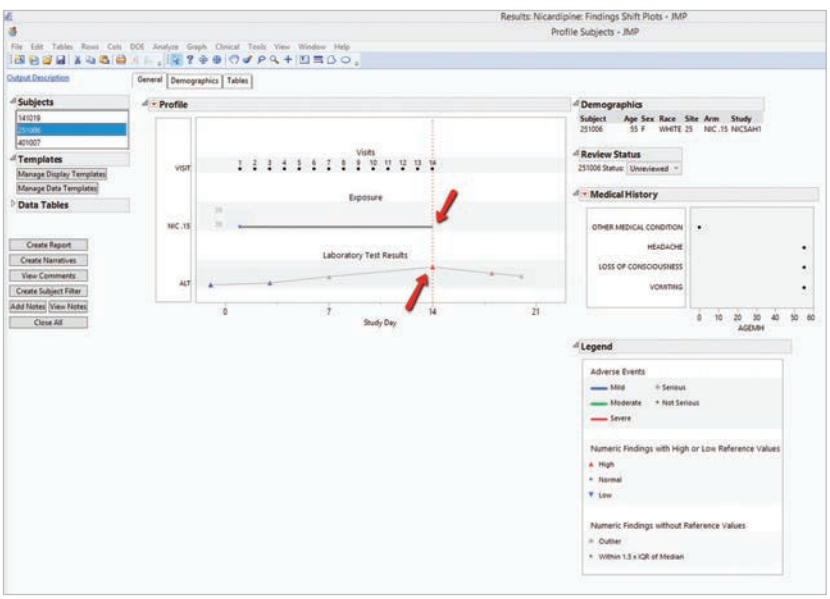


Figure 5: Conclude that the appropriate action with the study drug was taken using a subject profile

Every drug developer has a regulatory – and ethical – requirement to manage the risks and benefits of the drug under development.

Announcing JMP® 13

Lucky you. With the latest version of JMP, you'll spend less time assembling and preparing your data and more time uncovering the hidden possibilities your data holds.

Find out why luck is good, but the new features in JMP 13 are better.

1 Query Builder

The Query Builder for relational databases and SAS data sets has quickly become the preferred way to easily create accurate, repeatable and sharable queries – without having to write SQL code. However, until now you could not use Query Builder to join JMP tables already in memory, and nobody wants to create a database just to join data that is already in JMP. The JMP 13 Query Builder for JMP tables provides multi-table query and join capability through the same interface as the SQL Query Builder and SAS Query Builder. Once you start using Query Builder in JMP 13, you'll wonder how you ever lived without it. In addition, JMP 13 has many improvements that serve to streamline the data preparation workflow. Here are just a few:

- Run a script in a data table with a play button.
- Right-click "new column" options in the data table.
- Create and view formulas more easily with a retooled formula editor.
- Access the Local Data Filter in the main red triangle menu, rather than in the script submenu.
- Use SQL as a query language in JSL.

2 Dashboard Builder

Often, scientists and engineers must sell the change solution to others who may lack their level of analytics expertise. For clear and efficient communication, a single presentation-ready dashboard is the gold standard. In JMP 13, the Dashboard Builder aggregates JMP reports into a presentation-ready dashboard with only a few clicks and includes such time-saving features as:

- Drag-and-drop templates.
- Bigger, more visible drop zones, like Graph Builder's zones.
- One-click assignment of graphs as filtering devices.
- Real-time, customized analysis and layout reorganization.
- Tabbed dashboards made on the fly.
- An interactive HTML report that preserves the layout of your dashboard and can be shared with anyone – whether or not they have JMP.

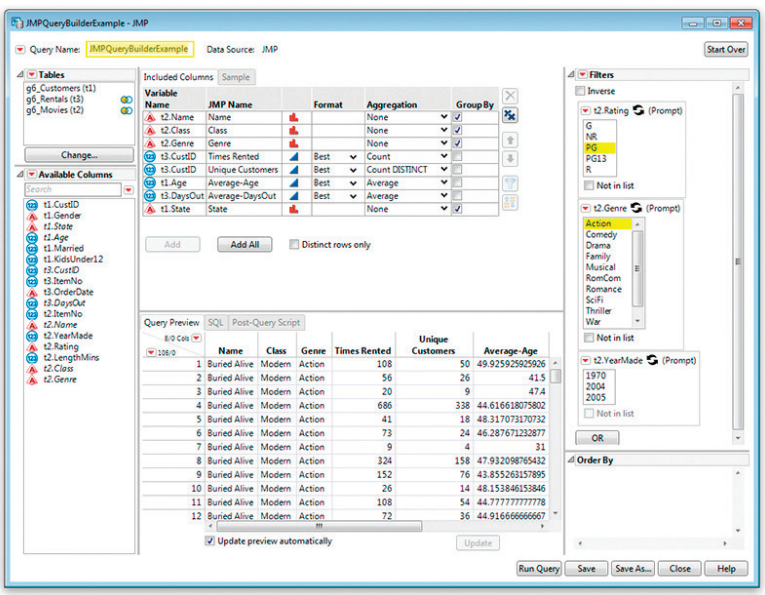


Figure 1: Query Builder

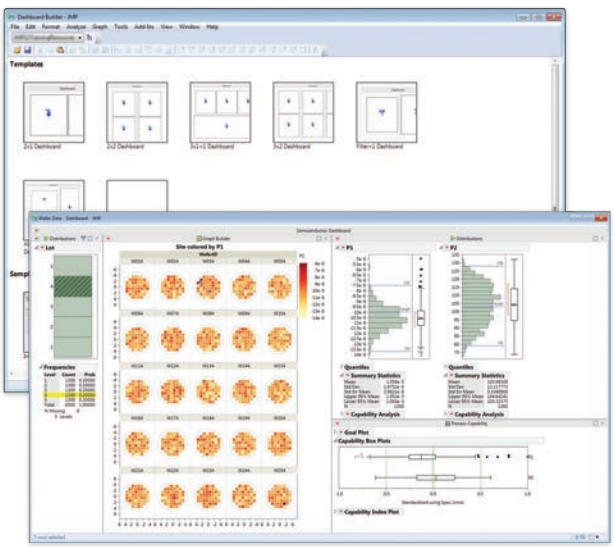


Figure 2: Dashboard Builder

3 Virtual Join

A virtual join eases the pain of joining and/or maintaining large (or many) tables; a physical join of a tall table to a wide one (for example, joining weekly summary values to data collected every second) can require more memory than is comfortably available. In the worst cases, a physical join was impossible. With the virtual join in JMP 13, you get the benefits of a join without creating a new table or using additional memory. The resulting memory savings can be significant – especially when columns contain large strings or images.

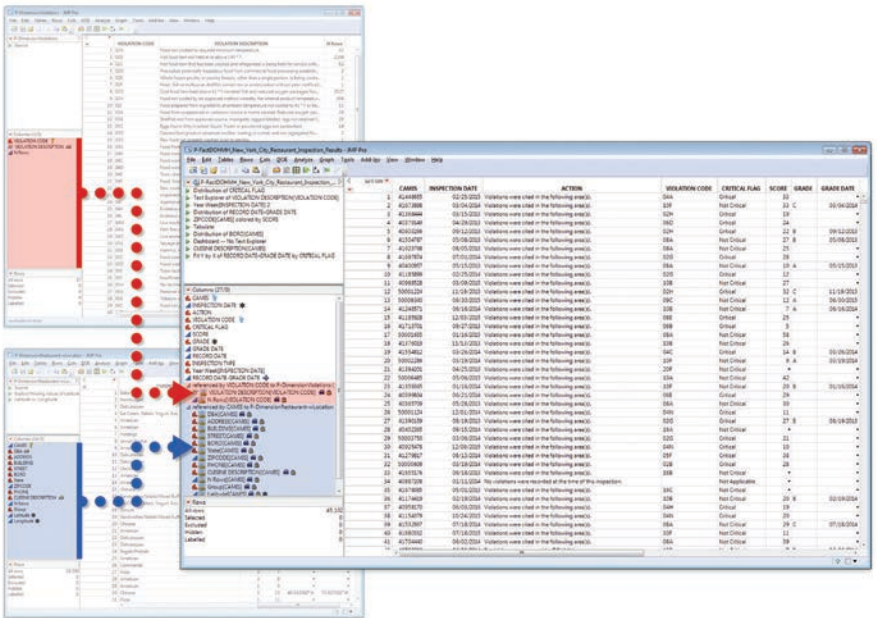


Figure 3: Virtual Join

4 Process Screening

In previous versions of JMP, there was no way to screen hundreds – let alone thousands – of processes automatically. As a result, process control analysts, Six Sigma practitioners, semiconductor engineers and other quality engineers had to manually triage these reports, which meant that most of the time allotted to searching for "problem" processes was consumed in viewing processes that were well under control. In JMP 13, there is a better way.

The Process Screening platform lets you quickly determine overall process health with the Process Performance Graph, and easily identify, using a customizable combination of tests and metrics, which processes warrant extra attention. It also lets you easily create control charts and process capability reports for any processes you choose.

With all the time you will save, you'll be able to focus more on problem solving, and less on firefighting.

The Process Screening platform in JMP 13 joins Predictor Screening (formally known as Screen Predictors) and Response Screening, as well as other modeling utilities (such as Explore Outliers

Figure 4: Process Screening

and Explore Missing Values) in the new Screening submenu of the Analyze menu.

5 Text Explorer

Free-text data exists in many forms: survey responses, repair logs, engineering reports and free-response fields are just a few examples. Text Explorer uses a bag of words approach to parse text into tokens to build a DTM. This allows market researchers, warranty engineers, medical professionals, engineers and scientists who have mountains of text-based data to easily triage and uncover the meaning in this data, rather than having to choose between processing it manually and ignoring it altogether. Text Explorer in JMP:

- Provides methods for basic keyword extraction.
- Has a local recode to clean up documents without altering the original text.
- Lets you find key words inside long text.
- Lets you process text data in informative ways.

6 Web Reports

Anyone who shares JMP results outside of JMP can do so more easily and effectively in JMP 13. Interactive HTML reports are great communication tools, but a single report usually cannot capture the entire analytic process. HTML Auto-Reports in JMP solved this problem by providing a linked, thumbnail-indexed collection of HTML reports – but installation of an add-in was required. The Create Web Report feature makes this functionality native to JMP, so no add-in is required. The list of reports includes a date/time stamp, comments and links to interactive HTML versions of the JMP report.

Also, the tedious and error-prone work of copying and pasting multiple JMP tables into Excel workbooks is now a thing of the past: the Create Excel Workbook feature allows users to create, update or append to an Excel workbook of multiple JMP tables – each in its own worksheet – with a single click.

7 Interactive HTML Reports

Graph Builder and Dashboard layouts are now supported as interactive HTML reports for communicating your findings with those who don't have JMP. You'll get interactive versions of Points, Smoothers, Ellipses, Lines, Bars, Areas, Box Plots, Histograms, Heatmaps, Mosaic Plots, Caption Boxes and Map Shapes. We've also worked to ensure your eye-catching dashboards built with the Dashboard Builder remain just as attractive when they are rendered as interactive HTML.

8 Graph Builder Enhancements

One of the most loved tools in JMP, Graph Builder, brings value to nearly every user, helping them make beautiful graphs to communicate their data's story as effortlessly as possible. In JMP 13, a host of new improvements and additions appear:

- Parallel Plots are now supported.
- Treemaps support an arbitrary number of nesting levels, with much greater control over the labels.

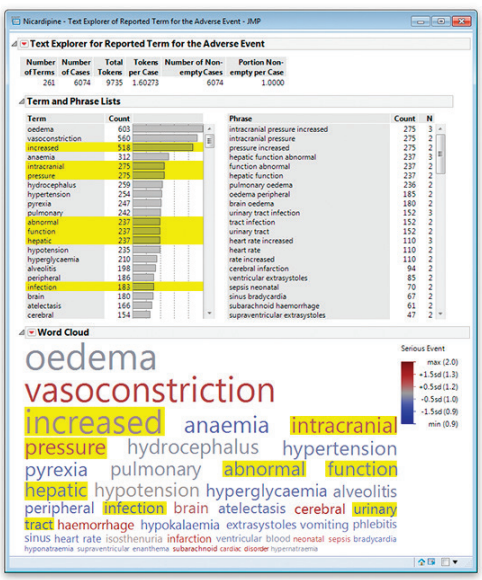


Figure 5: Text Explorer

Figure 6: Web Reports

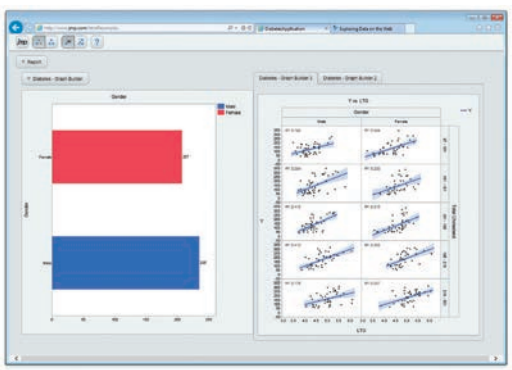


Figure 7: Interactive HTML Reports

- Alpha level can be specified for confidence bands.
- Ordering points is easier – just right-click the x-axis.
- Line element supports multiple "connect through" options and a centered step option.
- Legends, which have a new settings menu, can be placed inside the graph.
- Geometric mean has been added as a Summary option.
- The correlation coefficient can be shown for the ellipse element.
- The Response Axis property is now supported by all feasible elements.
- Multiple size and color variables are allowed.
- There are also some notable new features that make life easier not only in Graph Builder, but throughout JMP:
  - Quantile, Day of Week, Random Sampling and Moving Average one-click transforms.
  - Formula-based custom formats for numbers.
  - New color themes and lightness controls.

9 Compare Designs

In previous versions of JMP, it was difficult for DOE practitioners to compare designs easily, especially when there were more than two designs – it was a manual process requiring copy/paste or shuffling views among multiple design evaluations. Now, compare up to three competing designs and get all of the diagnostics in one report. Compare the designs with respect to power, correlation, prediction variance, aliasing and efficiency.

10 General DOE Improvements

With each new version, JMP furthers its world-class design of experiments capabilities. In JMP 13, key improvements include:

- Support for extra runs within Definitive Screening Designs (DSDs). These designs maintain the attractive aliasing properties of a DSD, while offering increased power due to the extra runs. Options for zero, four or eight extra runs are available.
- The Fit Definitive Screening Design platform, using a specifically tailored modeling technique for DSDs, has been added to the DOE menu.
- Simulate responses, which allows you to simulate data from a normal, binomial or Poisson distribution. Also, the DOE Simulate script in the generated design's data table lets you adjust simulation parameters after design creation.
- The DOE menu structure has been updated to provide more intuitive access, grouping the designers under descriptive headings.

11 Non-Normal Capability Analysis

In previous versions of JMP, quality assurance analysts, quality engineers, supplier quality engineers and Six Sigma professionals could perform capability analysis for normally distributed processes in the process capability platform, but for non-normal processes, they had to turn to the Distribution platform. In JMP 13, Capability reports can be created in a

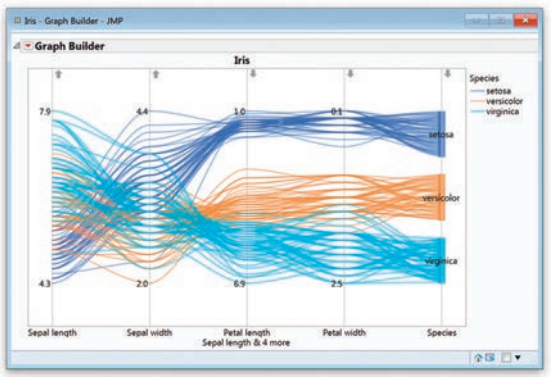


Figure 8: Graph Builder Enhancements

Figure 9: Compare Designs

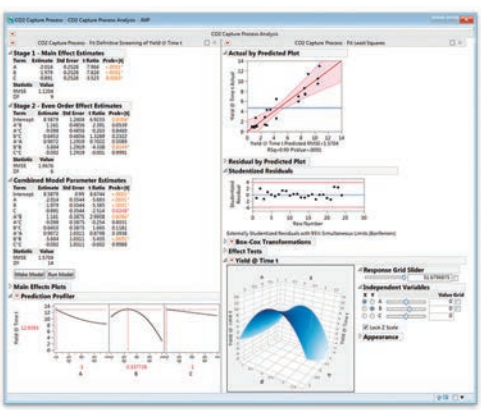


Figure 10: General DOE Improvements

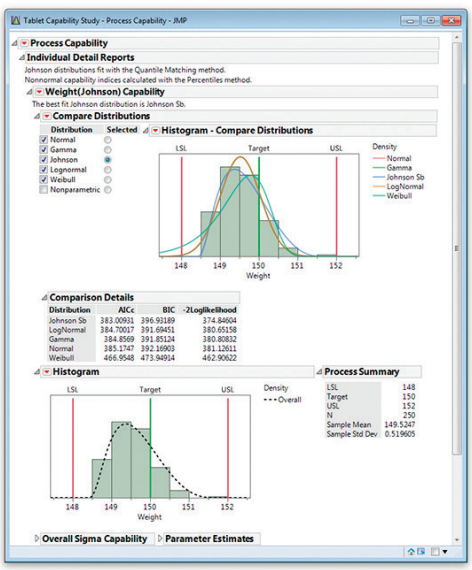


Figure 11: Non-Normal Capability Analysis

single platform, regardless of the processes' distributions. You can also:

- Independently set the distribution for each process in the report, and save these settings as column properties.
- Use nonparametric capability indices when the process distribution is non-normal.

12 Cumulative Damage Models

Design, manufacturing and test engineers need to be able to analyze time-to-event data under non-constant stress situations, but classical accelerated life tests (ALTs) assume constant stress over the life of any given unit. The Cumulative Damage platform in JMP 13:

- Provides an alternative to classical ALTs.
- Supports analysis of accelerated tests that specify varying stress profiles.
- Supports analysis of field data under arbitrary stress scenarios.
- Analyzes data having Step Stress, Ramp Stress, Sinusoid or Piecewise Ramp Stress profiles.

13 Multiple Systems Reliability Growth

The Reliability Growth platform offers the following new capabilities:

- Concurrent system analysis: Several similar systems are under study together. When any system experiences failure, the design correction applied to it is also applied to each of the other systems under study.
- Parallel system analysis: Several systems, which may or may not be identical, undergo testing in phases. A primary goal is to compare the reliability growth across systems.

14 Latent Class Analysis

Market researchers and social scientists need tools to reduce complexity in their data. Latent Class Analysis (LCA) is a statistical method for identifying unobserved class membership among subjects, using categorical observed variables. (For example, an insurer may wish to categorize people into different driver categories – the latent classes – based on the type of cars they drive, their accident records, and other publicly available data.) JMP 13 adds LCA to the newly reorganized cluster submenu, which includes a suite of row and column clustering techniques.

15 Multi-Dimensional Scaling (MDS)

MDS is a popular multivariate statistical technique used by sensory analysts, social scientists, market researchers and biologists to create a visual representation of the similarities among a set of objects. (For example, given a matrix of perceived similarities among various brands of cars, MDS can produce a map where points corresponding to similar cars are closer together and points corresponding to dissimilar cars are further apart.) There is little formal inference, such as hypothesis testing, associated with MDS, but it is a good tool for comparing products and making judgments about similarities without having to list product attributes,

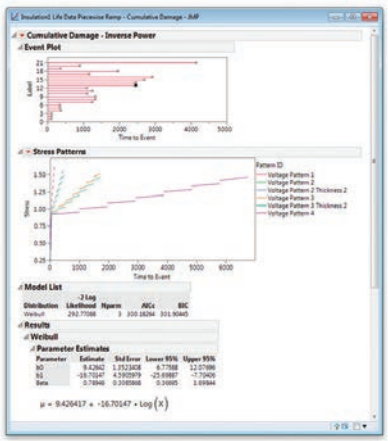


Figure 12: Cumulative Damage Models

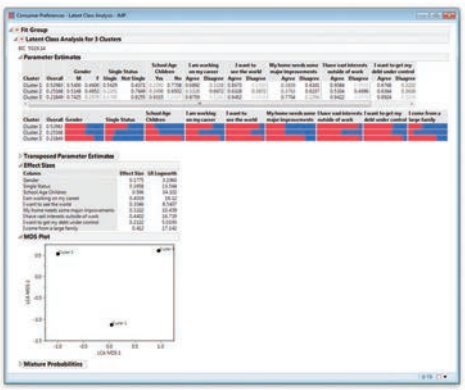


Figure 14: Latent Class Anaylsis

Figure 13: Multiple Systems Reliability Growth

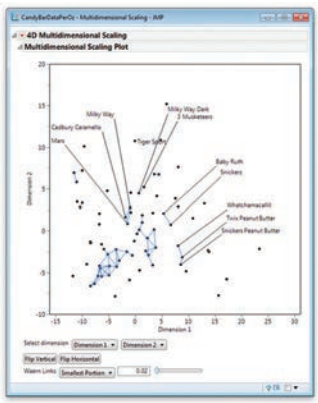


Figure 15: Multi-Dimensional Scaling (MDS)

as is required with other techniques, like conjoint analysis and factor analysis.

16 MaxDiff

MaxDiff analysis is a technique used to analyze customer preference. Similar to choice analysis, respondents simply select the best and worst choice from a set of fixed options. MaxDiff studies are often easier to execute than standard choice experiments, and in certain contexts, they can be more informative. With JMP 13, researchers can not only analyze MaxDiff studies, they can also design them using the DOE platform.

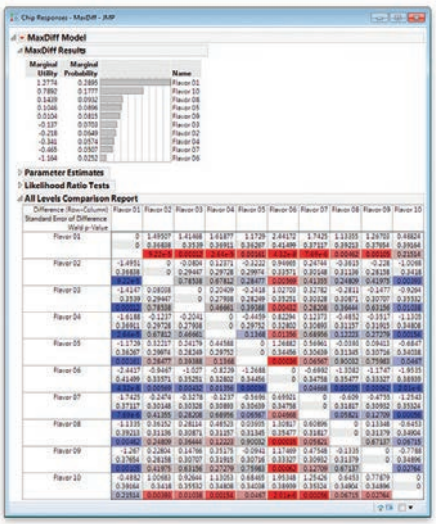


Figure 16: MaxDiff

17 No Choice Option in Choice Analysis

Many market researchers conduct choice experiments where “no preference” may be selected in lieu of the other options. Ignoring this possibility, where it exists, produces biased analysis results. JMP 13 now supports this important option.

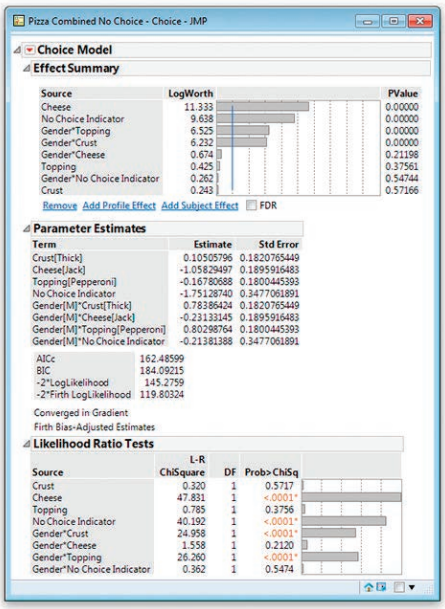


Figure 17: No Choice Option in Choice Analysis

18 JSL Improvements

For many users, scripting takes the everyday drudgery out of repetitive tasks, letting them focus on more interesting things: building models, designing and running experiments, and finding useful ways to communicate results. For others, scripting extends the JMP feature set, taking JMP beyond its native capabilities. New features and architectural changes in JSL include:

- Faster list processing.
- Auto-saving.
- Better namespace control.

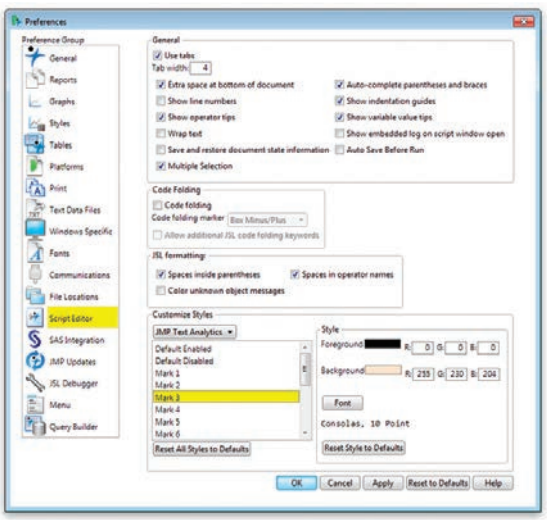


Figure 18: JSL Improvements

ONLINE

Want to learn more about what makes JMP 13 special? Visit our website: **jmp.com/jmp13**

Play with JMP 13 before purchasing with our free 30-day trial: **jmp.com/try**

Ready to get your hands on JMP 13? If you have a JMP annual license, your SAS administrator can give you access to the most recent version. If you don’t have JMP yet, get it online: **jmp.com/buy**

The power of Pro

As the advanced analytics version of our software, JMP® Pro contains everything users know and love about JMP® – and more

If you are ready to take your analysis to the next level, read on to learn about the new capabilities, new improvements to the predictive modeling workflow and performance improvements across most platforms in JMP Pro® 13.

1 Text Explorer Analytics

Introduced in JMP 13, Text Explorer is a platform for dealing with unstructured text data. While JMP 13 provides methods for basic word and phrase extraction, JMP Pro adds significant capabilities for multivariate analysis and dimension reduction, enabling users to incorporate text data into their predictive modeling tasks.

JMP Pro can take the types of text data that almost everyone has – repair logs, free-text surveys, long-form description fields and open comments, to name a few – and through the use of tailor-made analytics, convert that text data into numerical data that can be used directly by the modeling platforms of JMP Pro. Now you can take the information within your unstructured data and use it to enrich your favorite JMP Pro model, be that a Bootstrap Forest, Neural Net or Generalized Regression model.

By complementing the predictors you already have with the text data, you can build models with better external validity. You've already gone through the pain and expense of storing this unstructured data – why not put it to work for you?

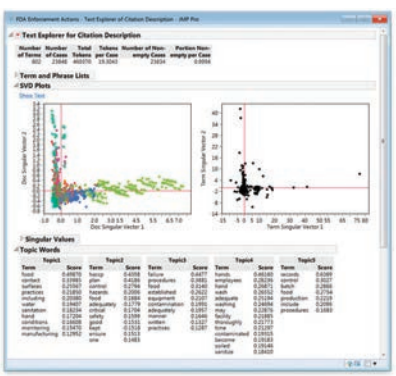


Figure 1: Text Explorer Analytics

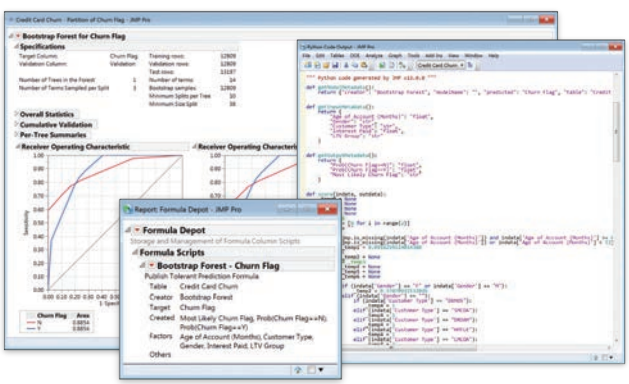


Figure 2: Formula Depot and Generate Scoring Code

2 Formula Depot and Generate Scoring Code

Formula Depot in JMP Pro 13 makes model deployment much easier, providing a central repository to organize, profile, compare and selectively deploy models in C, JavaScript, Python, SQL or SAS. Formula Depot allows you to store prediction columns in a central location, so your data tables stay uncluttered. It also allows you to easily compare any collection of models, each of which is informatively named so you know at a glance which platform created it.

3 Generalized Regression Improvements

The goal of Generalized Regression in JMP Pro is to satisfy all of your modeling needs. Generalized Regression is the modern approach of JMP to Generalized Linear Models and variable selection, as well as a world-class tool for analyzing DOEs and observational data – even life/survival data – all in one place. JMP 13 marks Generalized Regression's third release, and the tool continues its evolution with many additions and new features:

- A Double Lasso option: screens variables with an initial Adaptive Lasso pass, and uses the resulting parameter estimates as weights in a second Adaptive Lasso pass, further refining the model.
- Addition of a Two-Stage, Forward Selection option: An initial forward selection considers only main effects,

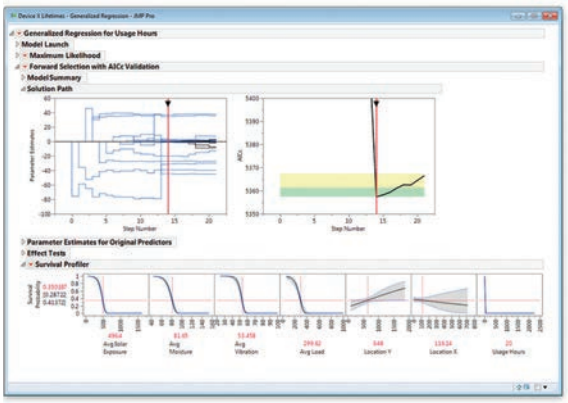


Figure 3: Generalized Regression Improvements

Figure 4: Repairable Systems Simulation

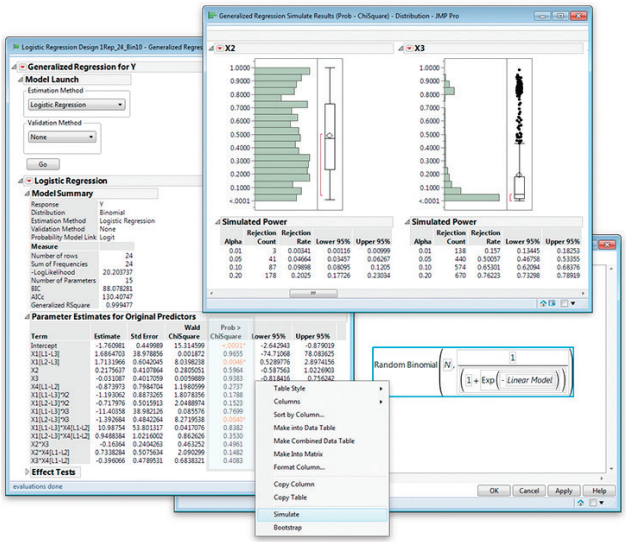


Figure 5: General Simulation Functionality

then a second pass considers interactions and higher-order terms. This technique has excellent properties for model selection in designed experiments.

- Handles censored data, allowing variable selection when fitting survival/reliability data; support for Cox Proportional Hazards; Support for Weibull, LogNormal and Normal distributions.
- General improvements:
  - New model selection criteria, ERIC, designed specifically for penalized regression problems.
  - ROC and lift curves.
  - Confusion matrices.
  - CDF and Quantile Profilers (like parametric survival).
  - Support for no intercept models and ordinal predictors.
  - Relaunch with active effects. Better model diagnostics. Ability to save simulation formulas that can be used in the Simulate utility.

4 Repairable Systems Simulation

A complex, expensive system built from many serviceable components – for example, an airplane engine – is usually repaired if possible, rather than discarded when it fails. With Repairable Systems Simulation (RSS), analysis of repairable systems is now possible, enabling you to answer questions such as:

- What is the mean time between failures over the "serviceable life" of the units?
- What is the availability of the units? That is, for what proportion of time are the units functioning?
- What are the expected repair costs over the serviceable lives of the units?
- Which components are most responsible for system downtime and maintenance?
- Which components should be opportunistically repaired/replaced while the system is down for another repair?

Since the RSS platform in JMP Pro 13 uses the same interface as the Reliability Block Diagram platform, reliability engineers won't have to learn a new interface – they can get right to work building models.

5 General Simulation Functionality

Statisticians and analysts use simulations to evaluate new statistical methods, estimate power for nonstandard statistical tests and perform parametric bootstrapping. In JMP Pro, access to general simulation capability is significantly broadened, and users no longer need to write custom JSL to perform the simulation and analyze the results. This capability is available from within most platforms having Auto Recalc or Bootstrap as red triangle menu options.

DOE is another area in which this simulation is useful. In designed experiments, the responses are not always distributed normally – or even approximately so. System tests in particular can generate data that is count – or pass/fail-based. Custom

Designer in JMP can now simulate realistic response data for these experiments, and JMP Pro users can use this simulated output, along with general simulation capabilities, to estimate the experimental power of the design.

Finally, Bagging (Bootstrap Aggregating) has been added to the Profiler in JMP Pro, making it possible for users to construct prediction intervals in settings where prediction interval formulas do not exist.

6 Hierarchical Bayes

Choice models can help companies decide which products and features are most important to customers, and how much customers are willing to pay for certain features.

A Choice model that treats all customers as if they were the same will produce estimates that average the preferences across all individuals. If, as is often the case, the individuals actually have significant differences in preferences, the "optimal" product as determined by the model may satisfy this "average" customer – who does not actually exist – but possibly appeal to none of the customers that do exist.

Fortunately, JMP Pro 13 supports Hierarchical Bayes, which can properly model these differences in preferences. Use it to get better results and higher-quality models from your Choice designs, maximizing the information gained from your limited run budget.

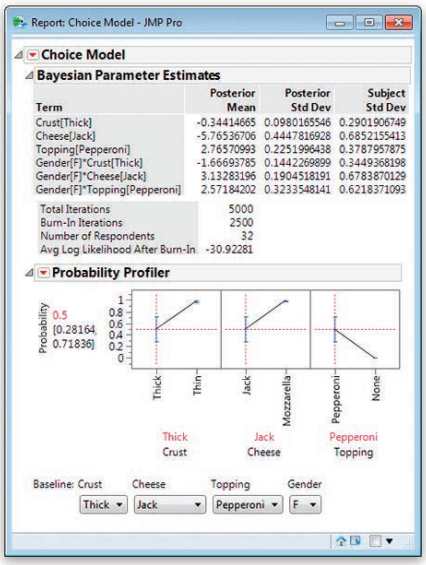


Figure 6: Hierarchical Bayes

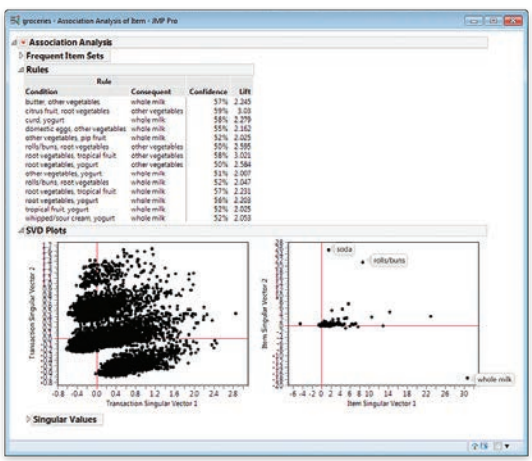


Figure 7: Association Analysis

7 Association Analysis

Association Analysis (often called Market Basket Analysis) is the identification of items that occur together in a given event, record or transaction. Here are some examples of association rules:

- Eighty percent of shoppers who buy product A also buy B.
- Forty percent of repairs involving part A also involve part B.

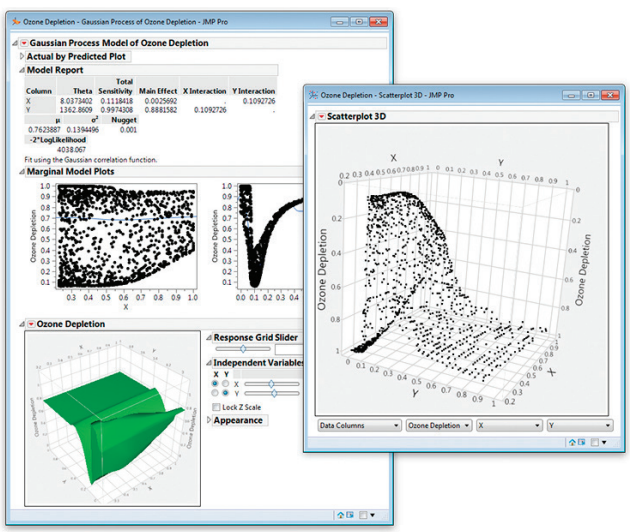


Figure 8: Gaussian Process

Figure 9: Mixed Models Improvements

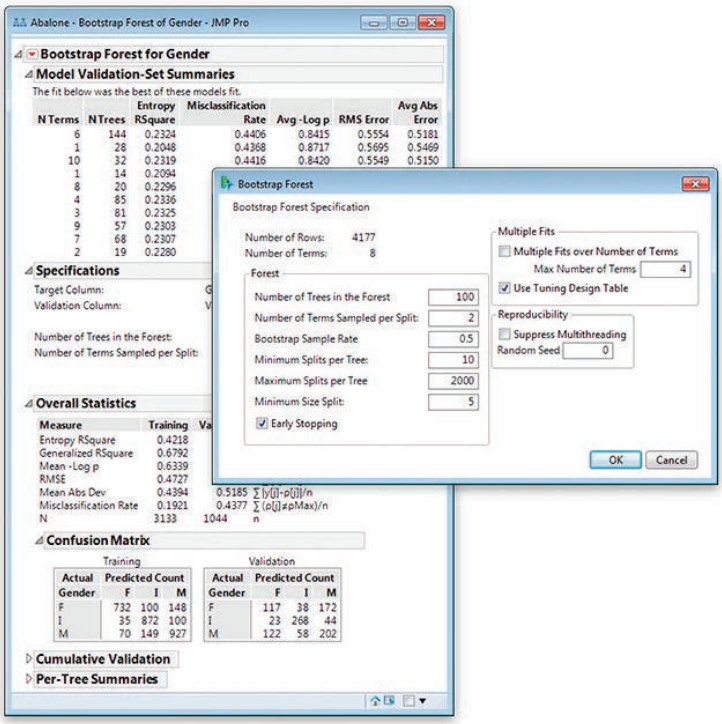


Figure 10: Partition Improvements and Naive Bayes

- Twenty percent of those having risk factors A, B and C have condition X by age 50.

Understanding these associations can inform decision making in a variety of contexts such as marketing, health care and product reliability.

With the Association Analysis platform in JMP Pro, you can also use Singular Value Decomposition (SVD), a dimension-reduction technique, to group similar transactions. The singular vectors can then be used in predictive modeling platforms.

8 Gaussian Process

Gaussian Process models are used to model the relationship between a continuous response and one or more predictors. In JMP Pro 13, Gaussian Process models run much more quickly on large data sets, and now support categorical factors.

9 Mixed Models Improvements

In JMP Pro, the Mixed Model personality of Fit Model offers several new covariance structures (Unequal Variances, Exchangeable, Antedependent, Toeplitz), greatly extending its versatility.

10 Partition Improvements and Naive Bayes

JMP Pro 13 gives you more control over the Partition platform: model tuning tables let you run models over a grid of parameter values, while the randomization in the Stochastic Gradient Boosting option for boosted trees helps protect against overfitting. The naive Bayes classifier is also available.

ONLINE

Learn more about the advanced analytics in JMP Pro 13: **jmp.com/pro**

Interested in playing with JMP Pro 13 before purchasing? Request a trial: **jmp.com/trial**

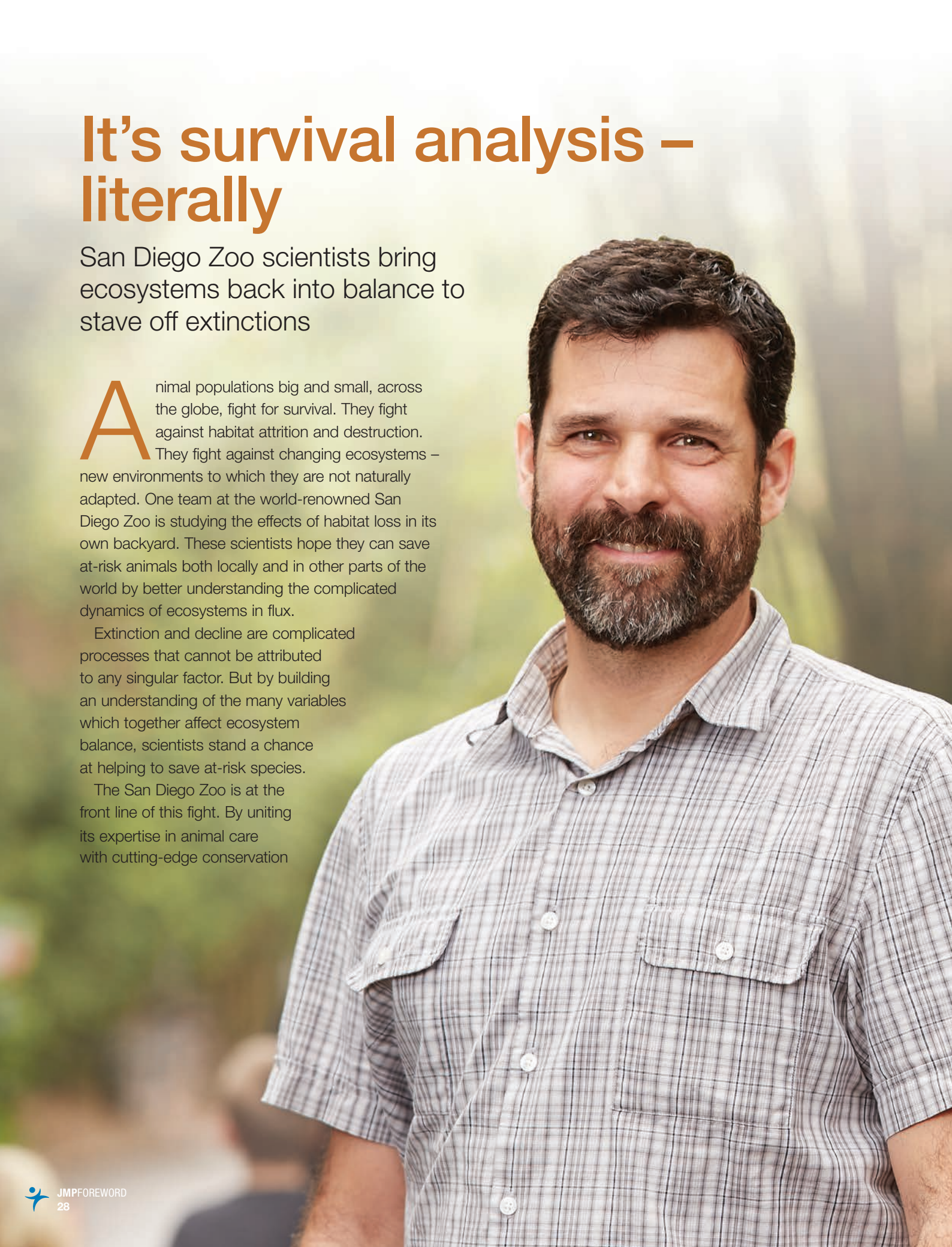
It's survival analysis – literally

San Diego Zoo scientists bring ecosystems back into balance to stave off extinctions

Animal populations big and small, across the globe, fight for survival. They fight against habitat attrition and destruction. They fight against changing ecosystems – new environments to which they are not naturally adapted. One team at the world-renowned San Diego Zoo is studying the effects of habitat loss in its own backyard. These scientists hope they can save at-risk animals both locally and in other parts of the world by better understanding the complicated dynamics of ecosystems in flux.

Extinction and decline are complicated processes that cannot be attributed to any singular factor. But by building an understanding of the many variables which together affect ecosystem balance, scientists stand a chance at helping to save at-risk species.

The San Diego Zoo is at the front line of this fight. By uniting its expertise in animal care with cutting-edge conservation



science, the zoo's researchers lead the global effort to end extinction.

An adaptive management approach to conservation aims at sustainable outcomes

Scientists of the San Diego Zoo's Applied Animal Ecology Division conduct important behavioral and ecological research to advance the institution's conservation mission. The division espouses an adaptive management approach – which means gathering information about the entire ecosystem over time in order to understand how endangered and threatened species are impacted by factors like habitat degradation, predation or competition.

In the case of San Diego County, California, scientists have noted significant habitat loss and fragmentation over the past few years due to human activity. Such changes have led to the diminution of native grasslands and a proliferation of aggressive invasive species.

Habitat loss in this region does not mean all native populations are thrown into peril; however, the reduction of one population can have a ripple effect that impacts the survival of other species. The California ground squirrel population, for example, has diminished markedly across the county as a result of habitat attrition. While the squirrel itself is not at risk locally, the western burrowing owl – for which the breeding population is indeed at risk of extirpation from San Diego County – relies on ground squirrels to provide burrows for nesting. When habitat loss pushes squirrels out, the western burrowing owl is at risk.

While a simple short-term solution to this problem might be to construct artificial burrows in which owls could nest, J.P. Montagne, Research Coordinator with the division, and his team see this option as unsustainable. The ultimate goal of conservation work, he says, is to help reestablish healthy, self-sustaining ecosystems that require little future intervention from humans.

"Typically, where there are ground squirrels, there is higher species diversity," says Montagne. In a way, squirrels are architects of the ecosystem. But when squirrel populations lack crucial components of the habitat they require for their own survival—the right soil, the right vegetation – they begin to decline. And so too do those bird and rodent species who rely on ground squirrels' cast-off burrows for nesting.

The theory goes that if you can bring back the squirrels, you'll also bring back the burrowing owl population. In reality, however, it's much more complicated than that. Squirrel populations' success is dependent on a wide array of environmental and behavioral factors that go far beyond a macro-view of habitat loss.

The San Diego Zoo seeks to understand those factors that facilitate the natural system's population dynamics. And that means experimentation and, ultimately, data analysis.

Strategic decisions grounded in data science

Together with his team, Montagne translocates squirrel populations to study variables affecting squirrels' ability to thrive, as well as the effects translocated populations have on the habitat itself.

Over the course of three years, the team has translocated over 700 squirrels. Prior to each move, Montagne performed a series of behavioral assays with the squirrels, a subset of which were radio-collared so that he could monitor them for three months following their release.

To complicate matters further, the team was forced to discontinue a handful of its relocation sites, supplementing the data by introducing new geographies with unique abiotic and biotic characteristics.

With a laundry list of temporal, ecological, geographic and what Montagne calls "animal personality"

> "Without JMP, I don't know that I could have mastered this kind of complex survival analysis as quickly and easily as I did. "
>
> J.P. MONTAGNE

variables, the resulting data set was a statistical challenge.

"It was a complex data set – we did translocations a few years in a row and had multiple plots at each site. And all these variables were really hard to wrap my head around," he says, "but JMP helped me figure it out."

Montagne runs multiple survival regression analyses on the number of days the squirrels survived as well as to evaluate movement parameters in the wake of a translocation. "Without JMP," he says, "I don't know that I could have mastered this kind of complex survival analysis as quickly and easily as I did. Its visualization and graphing features really make it easier for us as biologists to answer our research questions.

"Had I not had JMP, I would have had to learn R. I've used R in the past and it's also a great tool, but there's a huge learning curve – having to master scripting, for one. Most of us here at the institute are not statisticians. We're ecologists, biologists, reproductive physiologists. And we use statistics as a tool to help us answer the questions posed by our research."

Montagne says JMP cuts through the noise in his data to point at important relationships between variables. For example, analysis showed that the clay content of the soil was among the most important predictive abiotic factors of survival.

Data-driven conservation interventions help reestablish at-risk populations

But how does identifying predictive factors (and understanding their interrelationships) actually translate into promoting squirrel repopulation?

Montagne relies on the findings of his analysis to inform the strategic management decisions he and his team must make. By understanding the impact of predictive factors like soil quality, they'll be able to translocate squirrels to sites in San Diego County where they'll have the best chance at survival.

And with it, they hope to see the western burrowing owl begin to come back.

JMP has helped make sense of the complex data that drives this kind of conservation management decision making, Montagne says – even for scientists who, like his team, are less familiar with statistics. "Not only did I use JMP to run my analyses, you could say that I also used JMP to learn statistics," Montagne laughs.

JMP saves researchers a lot of time, he says – and the need to adapt research questions to what would otherwise be limited data sets. As a result, scientists now have more time to spend in the field playing an active role in the zoo's conservation efforts.

This intervention, like so many conservation strategies, isn't foolproof of course. But many reintroduction projects – take, for example, the high-profile Yellowstone wolf reintroduction efforts of the late 1990s and 2000s – have had tremendous success in recent years. And much of this success can be attributed to a turn toward systematic, scientifically informed management decision making.

ONLINE

For more information about the zoo and its conservation efforts, check out the website: **sandiegozoo.org**

Have an important problem to tackle? Interact and share knowledge with other JMP users: **community.jmp.com**

Want to see who else is making the world better? **jmp.com/success**

SAVE THE BIRDS!

The western snowy plover, a small, camouflaged shorebird that makes its home on the beaches of the Pacific Coast, frequently nested in the once widespread dunes. As did the California least tern. But due to habitat loss and increased predation, these birds have recently been added to the growing list of endangered and threatened species across North America.

Fortunately, however, the San Diego Zoo has taken up the case.

With an adaptive management strategy in their pocket and a collaboration with the US military (Naval Base Coronado and Marine Corps Base Camp Pendleton have joined forces with the zoo in working to protect native bird species), the Applied Animal Ecology Division now operates a robust bird research program in San Diego County.

Its mission? To gather data that could help inform conservation decision making. To better understand those habitat conditions that are necessary for nesting, a team of scientists tracked tern and plover survival and movement over several years. Collecting this data will allow them to answer a number of important questions:

- How long do these birds survive at a particular location?
- Do males and females exhibit different survival rates?
- Do birds remain year-round at breeding sites?
- Do chicks hatched at one site migrate to another to breed?

Even with an experienced field crew of scientists, however, such projects present many challenges.

"When we first started with our California least tern conservation project, it was hard to determine how many chicks were surviving," recalls Jeanette Boylan, Conservation Program Manager with the Applied Animal Ecology Division. "We knew they were hatching because we could see the chicks in the nest, but then they would leave the nest. And three weeks later we'd see fledglings and do a visual count.

"That method was relatively inaccurate, though. So we wanted to come up with a measure to more accurately estimate fledgling counts."

Survival and nesting data will now be used to create a complex meta-population model that will help guide researchers in making management decisions.

"We initiated a mark-recapture program for California least terns," Boylan says. "We were then able to take the data from this program and put it into JMP to produce survivability graphs. This strategy enables us to more accurately estimate when the chicks are dying.

"We can then take this information and determine which management practices are beneficial – or not beneficial – for this species. And hopefully we can modify our practices to increase chicks' survivability. Because if more chicks survive, the population will increase."

Support research conservancy at the San Diego Zoo: endextinction.org

Optimizing DOE for optimal product design

Innovation reigns at Gore Japan as engineers focus on waterproof material

Anyone who has ever hiked, biked, climbed mountains, or traversed the world's wilder places knows about GORE-TEX® products, the breathable, waterproof high-performance fabrics without which outdoor activities would be a little less enjoyable.

GORE-TEX® products, a flagship of W. L. Gore & Associates, are dynamic by definition. Not only are they designed to be physically flexible – an essential component of modern active wear – but they're also adaptable to different customer specifications. Water-wicking gloves for biathletes. Tactical shelters for the military. Waterproof jackets for all-weather hikers.

Gore Japan, a subsidiary of the US-based W. L. Gore & Associates,[1] works on the research and development, manufacture, and sales of GORE-TEX® technology. Gore Japan focuses on the expanded porous polytetrafluoroethylene (ePTFE) technology that makes this waterproof material – and Gore's other products – so innovative. The company's product lineup supplies a wide range of industries, from health care and semiconductors to aerospace, automotive and chemical. And if one thing can be

CHALLENGE

To improve process efficiency by adopting a statistics-driven operational approach across all levels of the organization.

SOLUTION

The company encouraged its researchers to use JMP to plan and implement efficient experimental design.

RESULTS

Gore Japan transformed its experimentation processes, more effectively and efficiently testing optimal combinations of factors affecting product quality. Employees' mastery of statistics is progressing, and those business areas in which JMP is involved have expanded.

[1] W. L. Gore & Associates, Co., Ltd. (Gore Japan) is the Japanese subsidiary of W. L. Gore & Associates, Inc., headquartered in Newark, Delaware.

> "Along with usability, another attractive feature in JMP is the ability to verify the validity of results visually."
>
> TATSUO YAMAMOTO

said about Gore Japan's multifaceted manufacturing processes, it's that data analysis has played a vital role in all aspects of the business.

A balance of advanced analytics and ease of use

Anticipating the challenges of growing demand for ePTFE technology, the company sought to integrate a new statistical analysis tool that was simple to use and produced shareable, easy-to-understand outputs.

"There are limits to what you can do with spreadsheets," says Tatsuo Yamamoto, a statistician at Gore Japan's Polymer Science Center, "so it was difficult for us to promulgate a data-driven mindset within our corporate culture without better tools. What we needed was a best-of-both-worlds solution fusing spreadsheets, which can only do simple things, and SAS, which requires programming skill."

In 2001, Gore Japan turned to JMP, a statistical analysis tool that has been the companywide standard ever since. And the number of JMP users continues to increase because JMP is easy to use, convenient and non-programming-reliant – in addition to aligning well within the scope of Gore Japan's work.

According to Yamamoto, JMP was already in use at Gore's US headquarters. Following tests using a variety of acquired data, JMP proved able to handle a wide range of statistical methods, from basic analysis to complex multivariate analysis, without relying on programming know-how.

"Along with usability and an abundance of functions, another attractive feature was the ability to visually verify the validity of analytic results," says Yamamoto. "With excellent profiling functions, in which multivariate models can be understood visually, JMP was the right choice for us."

Nowadays, almost all of Gore Japan's engineers use JMP to plan and implement efficient design of experiments (DOE). With it, engineers can easily run experiments that yield critical insight into the optimal combination of factors affecting product quality. And as Gore Japan's use of JMP has expanded over the years, so too have the software's capabilities.

Encouraging proficiency in analytics across the organization

In 2010, Gore Japan introduced an entry-level statistics training program, aimed at new recruits in technology with

little knowledge of statistical analysis. The program incorporates 10 practice exercises spanning four days – from univariate, bivariate and multivariate analysis to screening programs, response surface plans and verification tests.

Among the program's goals are to increase interest in statistics, introduce users to JMP basics and strengthen linkages between research staff and the company's statistics team.

"Achieving the balance between an experiment's reliability and efficiency is just the planning stage," says Yamamoto. "Because the reliability and reproducibility of results obtained according to advance planning varies, it is necessary to draw up designs for experiments taking a variety of factors into consideration… after specifying what it is you want to know. With the custom DOE capabilities of JMP – that allow for restrictions based on a consideration of multiple factors – innovative experiments can be flexibly implemented.

"In experiments that focus on conditions based on past results and experiences, you won't reach true optimum conditions. You have to determine the size of the range to search, gradually narrowing down the optimum and remaining aware of an optimization approach."

Visual capabilities enable hands-on learning

The statistical training of new Gore Japan recruits involves a design optimization exercise. By the end of the program, they'll be able to meet required specs with minimal variation. With the repeated process of prototype creation, experimentation and analysis, participants follow the same steps as actual business operations. In the process, they learn the importance of DOE, how to use statistics correctly and how to use JMP in operational processes.

Today's more thorough training has produced a new work cycle. Engineers consult with the statistics team before an experiment, planning and implementing DOE based on the team's advice. In the future, Gore Japan plans to enhance the business systems database with a similar approach.

And JMP will likely be used here, too, paving the way to still newer uses of data.

ONLINE

Technological innovation is at the heart of Gore's products: **gore.com**

Learn new strategies for DOE by watching our Mastering JMP series: **jmp.com/mastering**

Q&A WITH WILLIS A. JENSEN, GLOBAL STATISTICS TEAM LEADER AT W. L. GORE & ASSOCIATES

At W. L. Gore, the Global Statistics team is responsible for providing quality and statistical methodological leadership for the entire enterprise. The team has used JMP for data analysis (and more importantly, solving Gore's toughest practical problems) for more than 15 years.

Q: How does data analysis factor into your work at W. L. Gore?

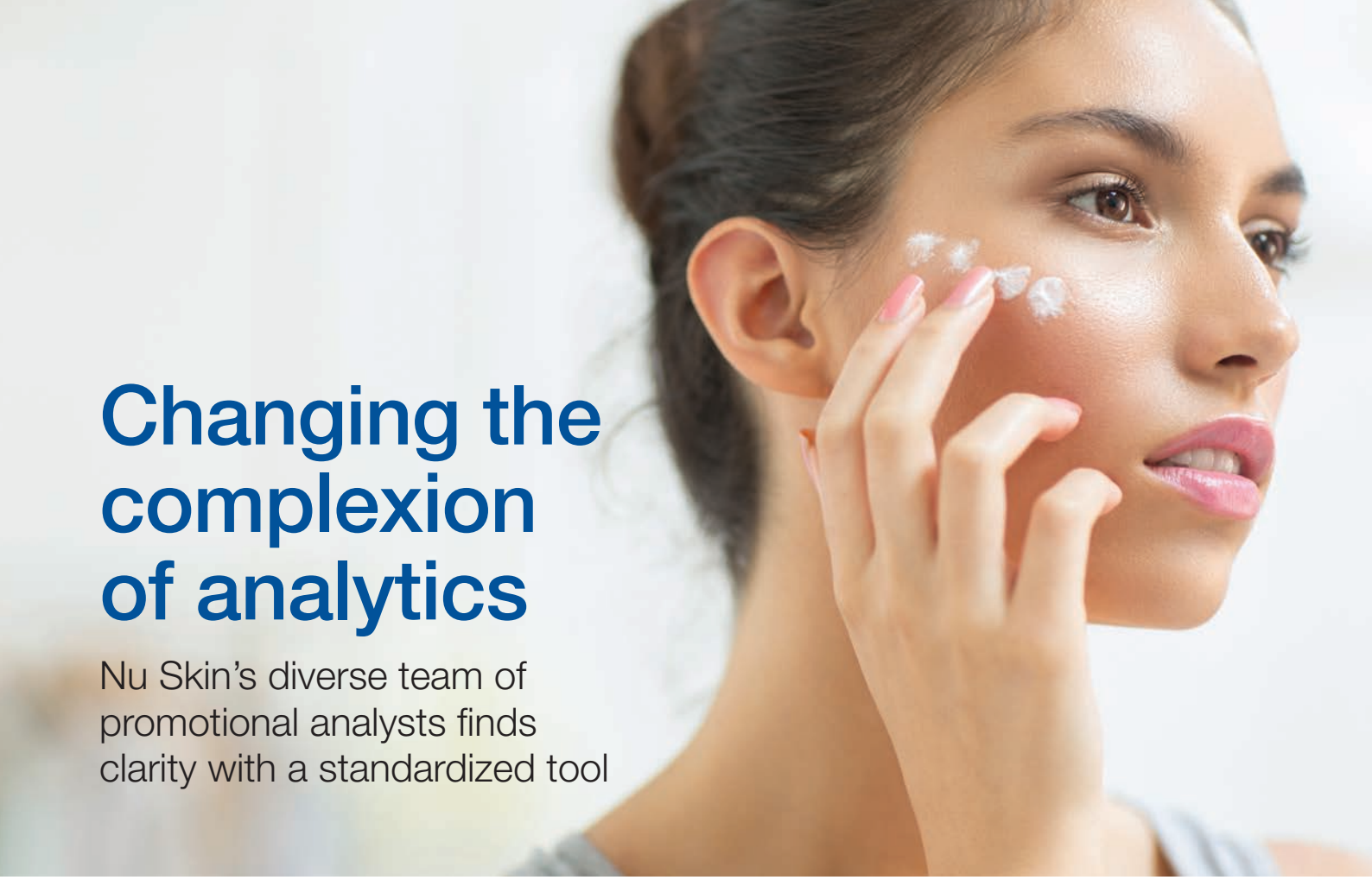
A: We work with such a wide variety of people throughout Gore, including manufacturing, engineering, finance, HR, marketing and more. JMP has such a wide variety of analysis tools that can be used across all of these areas. The DOE platform is our bread-and-butter tool that has been used in building many, many successful experiments throughout Gore.

Q: How has JMP helped you to overcome the data challenges you face? In other words, how does JMP make your life easier?

A: Graph Builder has become my go-to tool within JMP. It is powerful in its ability to quickly look at the data in many different ways. Whenever I show it to others who are not familiar with JMP, they are always amazed at how easy it is to create graphs and charts that communicate valuable insights from the data. Especially when compared to other tools at their disposal, such as Excel. In addition, large data sets with thousands or even millions of rows are no problem.

Q: Given the increasingly global nature both of markets and the business of manufacturing, how do you foresee the need for data analysis evolving?

A: As global markets become more prevalent, there is an increasing amount of connectivity between multiple data sources. This requires being able to quickly pull data from multiple databases and combine them easily prior to analysis. This is where some of the newer features of JMP are really valuable, specifically Query Builder and the column utilities to be able to quickly recode and clean up the data. JMP is definitely evolving to meet the data challenges of the future.



Changing the complexion of analytics

Nu Skin's diverse team of promotional analysts finds clarity with a standardized tool

Nu Skin develops and sells personal care products and dietary supplements in 54 markets worldwide. And its team of analysts working around the globe in these markets is representative of that diversity. To meet the sales challenges of a global operation, however, analysts needed to overcome barriers posed by language, culture, technical expertise and geography.

Brian Eastwood, a market analyst for Nu Skin, was charged with standardizing the company's global analytics to obtain a more accurate baseline on the impact of its promotional processes – no small task for a company whose international reach was so broad.

Developing a new companywide standard

Nu Skin's strategic goal? To improve service and growth in regions throughout the world by encouraging analytics-driven decision making. To support this focus on evidence-based business strategy, Nu Skin hoped to roll out an analytical toolkit with which regional groups might begin performing promotional analyses on their own.

Specifically, Eastwood sought to develop a tool to facilitate control group selection and analyze promotional uplift, partnering with Brigham Young University to develop an interactive JMP environment. With this approach, he hoped to leverage the user-friendliness and visualization capabilities of JMP to allow users to gain practical insights from even complex data sets.

Catering to all levels of analytics expertise and language ability

But with global operations come added challenges. Building a standardized tool for an international team meant Nu Skin had to take into account regional nuances, colloquialisms and, most of all, language barriers. After all, not all of Nu Skin's analysts worldwide are

CHALLENGE

Use promotional analytics to make fast, reliable decisions around the globe by standardizing analytics processes, in spite of the wide range of language and technical skills possessed by the company's regional analysts.

SOLUTION

Nu Skin developed an interactive JMP environment to facilitate control group selection and analyze promotional uplift.

RESULTS

Today, this JMP tool is used to perform product promotions analytics, thereby improving service and growth in regions throughout the world.

> "During training, we go through the process three times and the analysts can do it by themselves....The tool doesn't require English [language proficiency] – the buttons do all the work."
>
> BRIAN EASTWOOD

proficient English speakers. So while stakeholders made the decision to develop the tool in the company's official language, English, they sought to make it visually intuitive in such a way that textual labels would not be necessary.

"So long as Nu Skin's analysts learn the right sequence of buttons to click," Eastwood says, "they almost don't need to know the English words. The tool doesn't require English proficiency. The buttons do all the work."

A second challenge? Many regional analysts at Nu Skin had limited experience in data analytics, while some came to the team with graduate level education in a technical field. "We have quite the range of skills on our team," Eastwood explains, "but the thing that connects us all is an understanding of our business objectives and the right mindset. Still, we needed a tool that was more visual, and that's why we use JMP."

More accurate promotions analytics

Before Nu Skin developed its standardized tool, the company's promotions analytics suffered from a huge variance in quality due to the many different evaluation methods undertaken within the organization. Today, however, Nu Skin regional analysts have confidence in their visualizations and outputs. Their results are valid. So the findings they present to their bosses are also valid.

"The Nu Skin JMP tool was an absolute turning point for me," says Georgia Alevizopoulou, a Nu Skin regional analyst based in Belgium. "Using the JMP tool has greatly diminished the portion of my workload devoted to tedious standard analyses… and enables me to achieve a more reliable output. Moreover, the interactive plots have helped me a lot by providing a better visualization of the data, making it easier to communicate results or illustrate presentations.

We now have a more effective and efficient way of working, obtaining and sharing knowledge, and ensuring that a unified analytical approach is used across the regions."

Nu Skin is reaping other benefits as well. Today its analytics processes are simplified, faster and easier. Before the adoption of the new JMP tool, analysts only had time to run analyses once or twice a year at most. Now they're able to conduct promotional analyses once a week or even within a day.

Supporting an ongoing global partnership

The Global Analytics Council at Nu Skin meets quarterly to present and share recent promotional analyses. By doing so, the group reaps the benefit of candid feedback from colleagues with diverse perspectives. And regional analysts are now paired with a corporate analyst as a support partner. "These quarterly case studies help everybody up their game a little bit," says Eastwood. "So everyone continues to learn."

ONLINE

Need to use data visualization to communicate results? Here's how: **jmp.com/dataviz**

Nu Skin products are available in 54 markets worldwide. Find out more: **nuskin.com**

Data analysis helps track patients in the HIV care continuum

The Colorado Department of Public Health and Environment analyzes medical record data in efforts to link people living with HIV to treatment and reduce the risk of transmission

Ending AIDS as a public health threat by 2030 is feasible if high HIV-burden cities around the world fast-track their AIDS responses.

In the United States alone, there are currently 1.2 million people living with HIV/AIDS. And while only a few decades ago, an HIV/AIDS diagnosis was often fatal, HIV is no longer a death sentence. Today, a person diagnosed with HIV at age 20 – if started promptly on antiretroviral therapies – is expected to live a normal life span, with a highly preserved quality of life. Achieving viral suppression both improves individual health and greatly reduces the risk of transmitting the virus to a partner. Nationally, too few people living with HIV are retained in care or have achieved viral suppression; a problem disproportionately affecting youth.

The public health cost of going without care cannot be overstated: According to a 2015 study by the Centers for Disease Control, 92 percent of new HIV infections (as of 2009) were attributable to people living with HIV who were not in medical care.

For optimal public health outcomes, optimize the care continuum

Why are so few patients being served? As might be assumed from the statistics, many people living with HIV today face barriers to treatment; they drop off the HIV care continuum – that is, the "dashboard" model of care from diagnosis to viral suppression – at every stage. Lapses in care can be brought on by any number of co-occurring conditions including substance use and mental health disorders, homelessness, and chronic conditions, making engagement in HIV care and adherence to treatment more challenging. Physical barriers to care, such as access to transportation, still prohibit many from accessing medical services, particularly in rural areas.

In 2010, President Barack Obama announced the National HIV/AIDS Strategy (NHAS), making increased access to care for people living with HIV a major national priority. The Obama Administration's subsequent HIV Care Continuum Initiative, announced in 2013, called for increased work to address drop-offs in HIV care, underscoring the idea that to achieve optimal public health outcomes, more people living with HIV needed to remain in care.

In 2014, a joint United Nations Program on HIV/AIDS released new 90-90-90 treatment targets by 2020. The 90-90-90 targets are significant in that they are the first targets that explicitly focus on improving the HIV care continuum from diagnosis to viral suppression. Updated NHAS targets for treatment and care are:

1. Increase the percentage of people living with HIV who know their status to at least 90 percent.
2. Increase the percentage of newly-diagnosed persons linked to HIV medical care within one month of diagnosis to at least 85 percent.
3. Increase the percentage of persons with diagnosed HIV infection who are retained in HIV medical care to at least 90 percent.
4. Increase the percentage of persons with diagnosed HIV infection who are virally suppressed to at least 80 percent.

This targeted treatment strategy is fast being instituted by health departments across the country. The goal of this approach? To use HIV surveillance data to improve disease control interventions.

Colorado uses data analysis to reach out to its at-risk population

The State of Colorado is doing its part to help the 13,000 people in the state now living with HIV. In order to remain healthy, people with HIV need to obtain regular

"Data quality is a big issue for public health. And that black box… that's where JMP comes in."

ELAINE DANILOFF

medical care throughout their lifetimes. To get lapsed patients back into care, the Colorado Department of Public Health and Environment (CDPHE) is piecing together fractured electronic medical data from multiple providers and geographies. Doing so will enable them to identify and locate people living with HIV who need treatment.

For Elaine Daniloff, Epidemiologist and Data Scientist at the CDPHE, the fight against the HIV/AIDS epidemic is personal. She was working as a laboratory research assistant when AIDS first hit the United States in the 1980s and recalls seeing many people within the LGBT community – of which she is also a part – lose their lives to the virus.

Daniloff now works tirelessly through the CDPHE to help people living with HIV/AIDS. To do so, she relies on data. "My job," she says, "is to characterize the community of people living with HIV so that we increase access to care, improve health outcomes and reduce new HIV infections."

JMP® is 'indispensable' in ensuring the quality of electronic laboratory report data

Daniloff and other data scientists working in public health rely on data from electronic laboratory reporting (ELR) systems. And though ELR has become the mainstay of reducing the HIV/AIDS epidemic worldwide, Daniloff says, "there are gaps all the time – we didn't receive a report, for example, or we didn't know we received it. Or we have duplicates. Data quality is a big issue for public health. And that black box… that's where JMP comes in."

JMP, Daniloff says, "has been absolutely indispensable with data cleaning, quality control, reporting and visualization. No matter how difficult or time consuming, I always find a way using JMP to link records from different data systems, work with dates and calculate time elapsed for intervention outcomes."

Daniloff uses JMP to match patient data from the state's case-based surveillance system with three goals: the quality assessment of ELR lab tests, record linkage and deduplication. For record matching and deduplication, she says the Tables features in JMP are critical – Summary, Subset, Sort, Stack, Split, Join, Update, Concatenate and Missing Data Pattern. Daniloff also makes use of Database Open Table and Query Builder, and creates visualizations to communicate her findings to stakeholders at every stage of the care continuum.

After identifying individuals living with HIV/AIDS in Colorado who have fallen out of care, Daniloff alerts state social workers. Fortunately, she says, between the state, healthcare providers and insurance stakeholders, there are resources to help people get the medical therapies they need. And Daniloff's analyses don't end there: she also uses ELR data to study the UN treatment targets, and examines patient outcomes among those who transition between HIV care providers in Colorado.

"The public health focus on viral suppression is the most important measure because it encompasses antiretroviral therapy's treatment and preventative benefits," Daniloff says. "Viral suppression allows people living with HIV to remain healthy and significantly reduces (or almost eliminates) the chance of HIV transmission between partners and mother and children. In the future, I hope to use JMP to create models to estimate HIV transmission among different risk groups in Colorado."

ONLINE

Explore how the Colorado Department of Public Health and Environment is leading important statewide health initiatives: **colorado.gov/cdphe**

Learn more about the fight against HIV/AIDS: **aids.gov**

Read how others are using JMP to tackle interesting challenges big and small: **jmp.com/blog**

The Beijing Aquarium finds clarity in analytics

Water quality engineers increased tank visibility while decreasing energy use

Unlike their natural seaside counterparts, inland aquariums rely on advanced technology for life support and water treatment in order to sustain the sensitive aquatic conditions that will enable marine species to thrive even thousands of miles from the ocean. The Beijing Aquarium is a case in point. What was at the time of its construction in 1999 the world's largest inland marine facility, the Beijing Aquarium spans an impressive 120,000 square meters. With so large an aquarium at such great a distance from the sea's natural systems, water management presents some especially tough challenges.

Manufactured ecosystems present a puzzling cost/benefit quandary

Little is known, however, about causality in artificial aquatic life support systems. And there are few existing guidelines for managers tasked with optimizing water quality conditions in an aquarium's manufactured ecosystems, though with endangered species' survival in question, stakes are high.

The Beijing Aquarium is home to a pod of rare beluga whales. The beluga whale, according to the International Union for the Conservation of Nature, is

"Before we had JMP, we didn't have a good way to make accurate predictions. JMP has changed our perspective as to how we can better anticipate changes and fluctuations in water quality in the future."

LEI WANG

considered "near threatened" around the globe, though in Canada, shrinking population size in recent years has caused the government to reclassify belugas as endangered. Beluga whales are indigenous to the frigid waters of the Arctic Ocean, where their distinctive all-white phenotype, stocky body and heightened echolocation skills help them to survive extreme conditions.

While the Beijing Aquarium's belugas need not fear the natural threats of the wild – polar bears and killer whales are among their chief predators – their survival in captivity depends on the careful monitoring and management of their human stewards. Conditions in the beluga tanks should mirror that of the natural environment as closely as possible.

There is, however, just one hitch: high turbidity – a quality linked to animal activity – means low visibility. And low visibility means aquarium visitors won't be able to see the belugas. Moreover, maintaining constant high-visibility conditions is expensive and requires high energy consumption. Was there a solution to satisfy these seemingly competing interests? That task fell to the Beijing Aquarium's Lei Wang and his team of water quality engineers. And they turned to JMP.

"Before we had JMP," Wang recalls, "we didn't have a good way to make accurate predictions. JMP has changed our perspective as to how we can better understand and anticipate changes and fluctuations in water quality."

Engineers use models to identify beluga whale habitat parameters

Wang and his team ran a series of tests in JMP in which the output variable was turbidity and the input variables were time of day and circulation rate. They asked: What is the relationship between each factor and the output response? "Turbidity is linked to two factors," says Wang, "First is the recirculation rate of the water. Second is time of day, which is closely related to whales' biological clocks."

"With this model, we can simulate all kinds of situations so we will know what time of day to use what kind of circulation rate in order to provide sufficient turbidity in the tank. We decided that when there are visitors, we should turn up the circulation rate to decrease turbidity. But when it's closing time and the visitors have gone, we can decrease the circulation rate to save energy."

When the Beijing Aquarium deployed Wang's proposed strategy last year, administrators found that the results were very positive: A brighter display during the day provided a better customer experience while energy consumption actually decreased by two-thirds. The aquarium now uses only one-third the amount of energy it was using before, with savings of around $5,000 a month for the beluga whale tank alone – and that is one of its smaller tanks.

From trial-and-error to data-driven decision making

With the success of their data-driven management strategy for the beluga whale tank, Wang and his team put JMP to work investigating other areas in which the aquarium could improve its

processes. Until recently, ecosystem management decisions at the aquarium were being made based on precedent, past experiences and trial-and-error rather than on data. And the outcomes were not optimal.

What the aquarium needed was to systematize the decision-making process – observe results over time and develop standard guidelines. But with many workers having little to no technical training, the means by which the organization did this had to be accessible, says Wang. "Even after just a year and a half of using JMP, we made several meaningful discoveries with our data. And we learned that JMP was easy to use – even for people who don't know much about statistics."

In addition to determining correlations between water quality measures to balance the manufactured ecosystem, Wang and his team also started looking into the relationship between the aquarium's operation parameters and water quality measurements. They now generate a weekly report in JMP that serves as a visual representation of fluctuations in water quality over time. "Before we had JMP, we had to rely on a combination of Excel, SPSS and Orange to create these reports. But now JMP is the only software package we need," Wang says.

JMP® outputs facilitate regular reporting

Before the aquarium started using JMP, it would take Wang's entire six-person team a full week to produce the annual report. With JMP, it only takes one person two days to complete. "That's a huge time savings," says Wang. Moreover, "JMP provides an intuitive, interactive way to view outputs. The aquarium's high-level management doesn't have to possess much statistical knowledge in order to understand the implications of our work."

In the future, the aquarium plans to coordinate between departments so that all can collaborate on their efforts to understand how aquatic ecosystems function as a whole and, Wang hopes, use data to manage those systems more efficiently.

ONLINE

The aquarium is one of the newest additions to the prominent Beijing Zoo, which was founded during the late Qing dynasty in 1906: **www.bjzoo.com**

Learn innovative strategies for data mining in JMP: **jmp.com/data-mining**

JMP® IS AN ESSENTIAL TOOL FOR WATER QUALITY ENGINEERS NEAR AND FAR

Michelle Woolfolk, Water Resources Engineer for the City of Durham, NC, uses data analysis in JMP to monitor water quality and prevent eutrophication in lakes and rivers. JMP comes in handy, Woolfolk says, when her team is monitoring oxygen levels in lakes across the county. When oxygen gets too low, algae is usually the culprit. An overabundance of algae results in high levels of chlorophyll A, which reduces light penetration. And when algae dies and decomposes, oxygen levels drop, endangering the lake's fish.

Woolfolk, who keeps a steady eye on algae growth, is deluged with data. Routinely, 3,000 or more numbers lie behind a story that she can tell in one visual, easy-to-grasp JMP graph. "Would you rather look at 3,000 numbers, or look at this one picture?" Woolfolk asks. "Personally, I'd rather look at the picture. This work can get complicated, but with partitioning tools in JMP, you can illustrate your findings very clearly."

JMP offers a solid Partition platform that Woolfolk returns to regularly for answers to such questions as, "Is this what I hoped I would see? Or is it something else completely? And can I explain it?" The Partition platform creates a set of groupings of X values that best predict a Y value by searching all possible groupings. It selects the optimal options from a large number of possibilities, thereby reducing large problems into digestible pieces. Read the full story: **jmp.com/durham**

From scraps to interface in six weeks

JMP® helps to streamline a medical device manufacturing process with innovative Visual Six Sigma thinking

"Excellence is never an accident. It is always the result of high intention, sincere effort and intelligent execution; it represents the wise choice of many alternatives – choice, not chance, determines your destiny."

Laura Zambianchi cites this quote from Aristotle as a guiding principle in her team's work. Zambianchi is a Senior Filtering Material Specialist and Six Sigma Black Belt at Fresenius HemoCare, located in the northern Italian municipality of Mirandola. As you might expect, her work involves the extensive use of statistics, which she employs as a means of intelligent execution.

But she engages with statistics in a specific manner.

"I'm a physicist," Zambianchi says, "and the way I use statistics is not to confirm existing ideas but to formulate questions. The way I analyze data is recursive. I try, and try again, and then I modify. I find my way through by using data."

Inspired by theoretical physicist Ian Cox's groundbreaking book *Visual Six Sigma*, Zambianchi takes charge of the decision-making process by leveraging dynamic data visualization techniques. Turning directly to the data set for insight, she asks relevant questions and then makes wise choices based on her findings.

JMP software is designed to point scientists toward those findings.

Zambianchi and her Fresenius colleagues are frequently called on to create something entirely new. Very quickly. On a budget. Recently, they received an ambitious assignment: to help install and validate a new manufacturing process for filters used during leukocyte depletion, a critical step in blood transfusions.

This brand-new manufacturing model would incorporate new technologies. Her team's task was to develop an interface to manage and connect the tremendous amount of information the new system would collect about the effectiveness and efficiency of its operations, pinpointing inefficiencies as the team worked for continuous improvement.

In the end, Zambianchi's team delivered the finished plan on time in six weeks, and the line started to operate at target design efficiency two weeks ahead of schedule. "In six weeks, we went from scraps to the interface," she says. "JMP was the ideal tool. It's the only Six Sigma tool that's real-time and interactive."

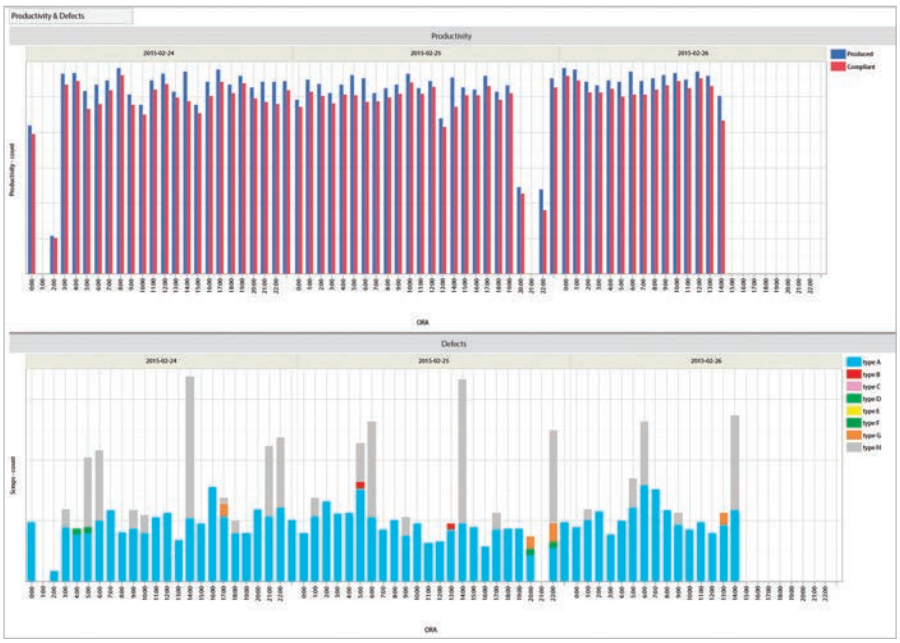
JMP Scripting Language (JSL) contributed to the success of the project, allowing the team to create a graphical interface for the process, with customized reports and graphs. JSL lets users write custom scripts to perform analyses and

visualizations not available in the point-and-click JMP interface, or to automate a series of commands.

"We weren't interested in the complexity of the scripting in every technical detail. We wanted a simple, lean tool that would allow for highly interactive data analysis. That's something that JMP does very, very well."

Real-world application

The Fresenius team (Sara Bergonzini, Emma Caiazzo, Luca Ferraresi, Alessandro Girotti and Zambianchi) comprises engineers and project managers with backgrounds in physics, informatics, material sciences, chemistry and mechanics. Though all team members had completed Six Sigma training, some had no previous experience with JMP software and its ability to link graphics and statistics to reveal trends, patterns and outliers in data.



Laura Zambianchi used JMP to create a streamlined, lean approach to examining productivity and defects.

After two days of JMP training with STATCON software consultants, the team dedicated some time to project scoping and design input, followed by two more days in workshops that allowed them to practice using the software.

STATCON's Sebastian Hoffmeister helped specify requirements and proposed a basic structure for a custom script in JMP Scripting Language (JSL). "This structure allowed the team to visualize all the relevant information about the production process just by plugging in the code, generating the required graphs and tables into the backbone code.

"JMP proved to be the perfect tool for the task," Hoffmeister added. "Having access to all the important methods and a large variety of data management routines made the development as straightforward as possible."

The Fresenius team then set about the task at hand. Team members met weekly to align the various process phases, then dispersed to conduct research individually. "The key to our success was working simultaneously but independently," says Zambianchi, a JMP user since 2011.

Team members used JMP to animate their reports with graphs and tables relevant to their assigned sections. The software allowed them to collectively produce a highly customizable report separated into sections representing each stage of the production line.

"The classical way of using statistics is to confirm an existing hypothesis," Zambianchi says. "But when you're in the shop, you're looking for real-time answers to questions that arise each day. What you want is something that works as fast as possible, something that's highly interactive and graphical.

"That's JMP and Visual Six Sigma. Every user, after very minimal training, can access the data, analyze it and get answers to their questions."

The team's final product is called a "process supervisor." It's inexpensive, simple and flexible, and allows for a complete, real-time picture of what's happening on the shop floor. It automates the analysis of all manufacturing steps, from component assembly to packaging, connecting to multiple data sources. That automation embeds not only the manufacturing phases but the quality inspections required for medical devices, and it allows the production line to operate with minimal operator oversight.

The quick turnaround provided concrete financial benefits, she adds: "When the line begins production, it initiates the return on the investment. So shortening that validation time makes a big difference. JMP allowed us to do that."

Improved production-line efficiency and fewer breakdowns offer additional savings. And when line operators receive feedback on performance throughout the process, they're able to provide proactive maintenance.

"The financial benefit of this project is essential to the automation approach," says Fresenius Hemocare Italia General Manager Alberto Bortoli. "High speed and reduced costs commonly associated with automation in the manufacturing

> "I'm a physicist and the way I use statistics is not to confirm existing ideas but to formulate questions."
>
> LAURA ZAMBIANCHI

environment often sacrifice quality. But not in this case – we upgraded the information flow and reaction time to ensure our product quality standards."

On the train

In every manufacturing process, Zambianchi explains, you have a physical flow of items – the components that are assembled in the various phases of the process.

"But you must keep in mind that it's not only the components that are flowing throughout the process," she says, "but also the information."

"For example," she explains, "there's the flow of quality control information throughout the process. You can have defects at almost every station, and if it's a medical device, there must be traceability for each component that you produce."

The job of Operational Excellence is to detect mistakes or, even better, to prevent them before they occur. Bergonzini suggests visualizing a subway system: "When a component is traveling from station A to station B, all the process features, the X's, and all the outcome features, the Y's, are traveling with it.

"Every process phase is like a station, and every station is collecting data in its own data set with its own format. There are stations, for example, that are collecting data on a time-series basis and other stations that are collecting on single-item or batch traceability."

Only through the proper process can the desired destination be reached. And to produce an efficient root-cause analysis, all this data must be joined, to create in real time what in Six Sigma jargon is called a transfer function.

"JMP allowed us to analyze, in a very interactive and visual way, what had an impact, or didn't, on that process," Zambianchi says.

A modular approach to programming challenges

"JMP unlocks huge potential," Zambianchi says, citing JSL and the ability to manage large volumes of data interactively as the key features employed in this project.

JSL was particularly attractive, she says, in that it offered a debugger, which helps identify the point at which a script causes an error or fails.

Zambianchi says the script they produced was used to reduce time to market, and is still used to improve process yield, conduct root-cause analysis, track material consumption, track production costs, perform preventive maintenance and more.

"JMP played a big role in making this possible," Zambianchi says. "It enabled us to approach complex programming projects modularly. The feasibility of the end result then becomes only a matter of resources and time."

True to Aristotle, the successful outcome was no accident. Excellence prevailed, a result of intention, effort, execution and wise choices – including the right software tool for the job.

The Fresenius team (left to right): Laura Zambianchi, Alessandro Girotti, Sebastian Hoffmeister (STATCON), Sara Bergonzini, Luca Ferraresi and Emma Caiazzo.

ONLINE

Learn about JMP for quality engineering: **jmp.com/quality**

Read a chapter from *Visual Six Sigma*: **jmp.com/vss**

Fresenius has more than 200,000 employees in more than 100 countries. Learn more: **fresenius.com**

Pioneering scientific analysis

An exploration geologist develops a new method for geostatistical cluster analysis

For untold centuries, mankind has scoured the earth for gold and other precious metals. Some with skill. Some with luck. All with the tools of their times. And all with a spirit of discovery.

One of today's foremost exploration geologists is Steffen Brammer, Senior Resource Geologist at Perseus Mining and independent consultant for Brazca. While his excursions into the field are less arduous than those of his predecessors, his goal is the same: To determine whether an ore deposit is of sufficient quality to warrant the expense of a full-fledged mining operation.

Brammer has spent most of his professional life working in mineral exploration and extraction in Africa. First in Ivory Coast and other parts of the West African region, then in Lubumbashi, a city in the southern part of the Democratic Republic of the Congo. And since 2008, Ghana.

"Twenty years ago, we were always on the road," Brammer recalls. "We'd pack up and just drive down a dusty lane armed with a compass and a satellite phone that we'd use once a week just to let people know we were still alive."

An unusually complicated data set necessitates innovation

But with the advent of new digital technology, mining industry expectations shifted. Exploration geologists like Brammer could now more readily identify deposits, but accessing these often complex bodies presented far greater challenges.

"We made a real discovery in Ivory Coast," Brammer says. "An ore body there exhibited two types of gold mineralization: One was moderately graded and the other

was small, but full of gold nuggets. We needed a new method with which to split a bimodal data set. Otherwise, these two different ore shoots would have to have been analyzed together – resulting inevitably in what would have been an inaccurate estimate."

To complicate matters further, Brammer faced a tight deadline: The company needed to go into production within the year.

From field exploration to digital discovery

Fortunately, Brammer is well equipped to handle this sort of challenge.

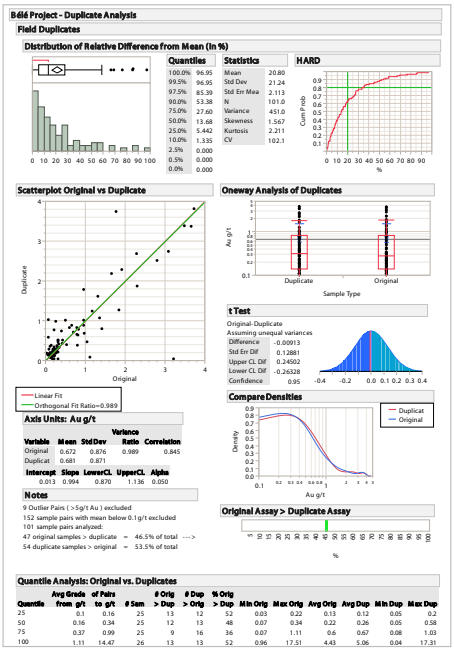
"In 2005, I was working in the Congo's Copperbelt region. I was sitting in the mine, contemplating how best to do my quality control calculations. But I had no software. All I had gotten from the company was a geological package. Not even Excel.

"When I first came across JMP, I bought it on my own dime. I began using it to run charts and plots – but soon realized that even the table options and summaries were a huge help. Today, I use JMP along with a geology software package like AutoCAD. I've written scripts in JMP to get beautiful reports, charts, graphs and analysis."

Brammer says the JMP application he developed with JMP Scripting Language (JSL) is absolutely invaluable in the face of challenging data sets – like the complex bimodal data from his days in Ivory Coast.

A new method for 'high-nugget' gold deposits

Applying geostatistical cluster analysis to the decomposition of mixed data with spatial information isn't a novel practice. This approach is not routinely applied, however, as cluster analysis isn't available with any of the usual geological software packages, making it especially difficult for non-statisticians. Various methods do exist for this purpose; however, where individual clusters are intertwined with irregular, discontinuous or complex geometries, conventional methods struggle or fail.



After an initial estimate of the statistical moments of the underlying components, a series of search trees are built through the sample grid, and samples are allocated to one of the conceptual target populations, depending on their probability density functions. Thus the mixed data set is split into its components while maintaining the spatial relationship within and across individual clusters.

Brammer's new approach was developed in JMP and implemented exclusively with JSL. After an initial estimate of the underlying components' statistical moments, a series of search trees are built through the sample grid. Samples are then allocated to one of the conceptual target populations, depending on their probability density functions. This way, a mixed data set can be split into its components while maintaining the spatial relationship within and across individual clusters.

The goal of this method is for the mining industry to unravel the phases of multistage mineralizing events among complex ore bodies. However, it is applicable to virtually all disciplines in the natural sciences, including environmental science, hydrology, biology and agriculture. The method also extends to every discipline where the spatial position of data matters, such as image processing, logistics, marketing and pattern recognition.

There are four benefits of Brammer's method:

1. It is computationally inexpensive.
2. It can be completed without expert assistance, and is thus suitable for universal and routine everyday use.
3. The user maintains editorial control over the final product because the results are not presented as categorical sets of clusters, but rather in the form of probability maps that acknowledge the possibilities of misallocation.
4. The method is easily applicable with commercial off-the-shelf statistical software (such as JMP).

"I've become a diehard fan when it comes to JMP," Brammer says. "I'm an exploration geologist – we're curious about discovery, and so is JMP."

ONLINE

Use JSL to create customized analytical applications and dashboards: **jmp.com/apps**

Read about interesting applications of JMP in many other industries: **jmp.com/success**

Visit Brazca online: **brazca.com**

Data-driven Quality by Design boosts drug development

Mitsubishi Tanabe Pharma improves processes for bringing new drugs to market

From basic research through nonclinical testing, clinical trials, approval and production, it takes anywhere from nine to 17 years to bring a new drug to market. And that's only if the drug makes the cut. In fact, only about one in 5,000 new drugs ever makes it to a pharmacy.

Odds like those can drive drug prices sky high, making them costly for cash-strapped consumers whose health depends on access to essential pharmaceutical therapies. And keeping prices down can be tricky. So the Mitsubishi Tanabe Pharma Corporation (MTPC) seeks to optimize quality at all stages of the production process. By increasing quality and efficiency, MTPC can offer state-of-the-art drugs that aren't cost-prohibitive.

Fortunately for MTPC, it's ahead of the game in systematizing process improvement decisions thanks to its analytics-driven approach.

The limited nature of spreadsheets left analysts wanting a higher-power solution

MTPC is a Japanese pharmaceutical company known for the arthritis remedy Remicade and the allergosis treatment Talion. These drugs, along with all of MTPC's products, were the outcome of a lengthy research and development process.

"We study overall formulation design, including pharmaceutical formulations and production methods, and conduct research to optimize multiple quality characteristics such as the hardness, disintegration properties and dissolution properties of tablets," says Keisuke Takagaki of MTPC's Pharmaceutical Research Labs.

Since the company's early days, however, MTPC's Pharmaceutical Research Labs have redesigned its approach to research. In the past, MTPC labs made minor changes to standard pharmaceutical formulations and production methods using observed, empirical judgments in order to optimize them for the desired quality characteristics.

"We were studying the process parameters and formulation factors affecting pharmaceutical quality based on

CHALLENGE

Gain insights into overall formulation design that will help to optimize multiple pharmaceutical quality characteristics.

SOLUTION

MTPC researchers used JMP outputs to make pharmaceutical development processes more efficient and reliable.

RESULTS

By using JMP to calculate processing parameters, MTPC was able to both optimize variations in quality characteristics and study design space settings for the dissolution rate of a new drug based on QbD.

"We look to JMP to help guide us through experimental design and, from the data we acquire, we use JMP to derive values for each processing parameter, thereby optimizing variations in quality characteristics."

KEISUKE TAKAGAKI

experimental design," Takagaki explains. "We recognized that there were many analytical operations that could help us to optimize these factors."

The tool used for this task? Spreadsheets. "Unfortunately spreadsheets were very limited in what they could offer. That's why we started looking to introduce a new statistical analysis tool." And most importantly, notes Takagaki, that tool had to be easily operated by researchers otherwise unfamiliar with statistics and programming. So the MTPC Pharmaceutical Research Labs turned to JMP.

Analysis helps pinpoint the factors that most affect product quality

The Pharmaceutical Research Labs now use JMP to design more streamlined systems capable of developing and producing pharmaceuticals that meet MTPC's high quality standards. In recent years, formulation designs that correlate with effectiveness and safety – in addition to formulations and production processes that can stand up to fluctuations in ingredients and process parameters – are a major focus for MTPC.

Equipped with a new tool, Takagaki's team faced their first task: to investigate how the dissolution rates of MTPC's current product offering might be improved. JMP helped researchers to search for the cause of trend abnormalities occurring in tablet dissolution rates, an important quality characteristic for all medications. With univariate analysis, researchers isolated the factors affecting drug dissolution rates and investigated the multivariate correlation between variables. All told, four factors were identified as affecting dissolution rate.

"Through multiple linear regression analysis, we determined that there was an especially strong correlation between the granules' mean particle diameter and tableting pressure," says Takagaki. Due to government regulations and standard tablet hardness and thickness specifications, however, it is difficult for researchers to make modifications to tableting pressure.

"Instead," Takagaki says, "we focused on the granules' mean particle diameter and, after using JMP to analyze the dissolution rate and relationship between factors, reworked our production methods to decrease the number of coarse particles and increase the number of fine ones, thereby improving the dissolution rate overall."

Taking action to advance quality and reliability

Researchers have recently implemented pharmaceutical development and quality risk management strategies based on concepts proposed in the ICH Q8, Q9 and Q10 guidelines, including Quality by Design (QbD) and Design Space (DS).

"For us here in the lab, QbD means we're setting each production parameter for the mixing and tableting processes involved in producing an orally disintegrating tablet," says Takagaki. "We look to JMP to help guide us through experimental design and, from the data we acquire, we use JMP to derive values for each processing parameter, thereby optimizing variations in quality characteristics."

Takagaki and his team have also studied the DS settings for the dissolution rate of a new drug they're developing – a conventional tablet. Specifically, they used an initial risk assessment and analysis of variance to narrow down the major factors affecting dissolution rate. Finally, they've also designed a response surface plan for two selected factors, the drug substance particle diameter and the granule particle diameter, and measured dissolution rates.

Going forward, MTPC's Pharmaceutical Research Labs will continue to use statistical methods to deepen the company's problem-solving abilities. "In the future, a statistical approach to pharmaceutical development and research will be essential," says Takagaki.

ONLINE

Go beyond spreadsheet analytics with visual data discovery: **jmp.com/beyond**

Make your Quality by Design Lean: **jmp.com/qbd**

Learn more about global, research-driven Mitsubishi Tanabe Pharma: **http://www.mt-pharma.co.jp/**

Closing the skills gap

Cal Poly develops the next generation of statistics professionals

What do you want to be when you grow up?

This innocent childhood question is one we've all been asked. And have answered: Truck driver. Teacher. Artist. Astronaut.

How old were you when you decided and knew it was the right answer? 23, 24, 25 …? Maybe you're still figuring it out.

Today's college students often know, upon arrival, how they want to use their degrees in the real world. At California Polytechnic State University, or Cal Poly, students are required to declare a major as freshmen.

Karen McGaughey, Associate Professor of Statistics, says they show up ready. "Today's Cal Poly students have had more exposure than we did as kids. They're pretty sure statistics is what they want to do."

Half of the university's statistics majors will go on to graduate school; most of the others go directly into the workforce. In the fall of 2015, the statistics department was expecting its largest incoming class – 35 students. Two years earlier, McGaughey and her colleagues had welcomed just 15 freshmen.

According to McGaughey, this growth is a good start to fill the world's growing need for statisticians. "There's a need for statistics, regardless of where you end up, just to make sense of our world," says McGaughey. Like Cal Poly, universities across the United States – and indeed, around the world – are working to establish statistics departments, broadening the discipline from its former home as a small subculture within the field of mathematics.

Hollywood makes statistics sexy

Pop culture is now brimming with instances of statistics in books, movies, sports and music. The internet provides infinite access. Students can literally search online for any statistical data they want. "Movies like *21* and *Moneyball* let students see statistics at work," McGaughey adds. "Then they all think, 'I want to work in sports for the San Francisco Giants.' Of course, they quickly see other applications of statistics, too. Students are surrounded by statistics, just generally in the world. Statistics is everywhere. Still, statistical literacy isn't where it should be."

McGaughey is a grader for the AP Statistics exam in Kansas City, MO, where AP Stats got its start in 1997. That first year, there were 7,000 exams. By 2015, there were 200,000. This growth is a good start to preparing students to fill the enormous and increasing demand for statisticians and data analysts trained to extract meaningful information from overwhelming volumes of data.

"I see that statistics is becoming part of the common core education, teaching kids how to think. Specialty K-12 schools that focus in science, technology, engineering and math (STEM) have opened doors earlier for students interested in statistics," says McGaughey.

A global shortage of analytically minded jobseekers

With the rise of data-driven decision making in recent years – be it in industry, government or business – the demand for statistics-savvy job seekers is fast

"There's a need for statistics, regardless of where you end up, just to make sense of our world."

KAREN MCGAUGHEY

increasing. Research by the McKinsey Global Institute forecasts that by 2018, the United States alone could face a shortfall of 140,000 to 190,000 people with deep analytical skills. The study also projects a deficit of 1.5 million managers and analysts with the know-how to use the analysis of big data to make effective decisions.

As the McKinsey report indicates, leaders in every sector, not just a few data-oriented managers, will have to grapple with the implications of big data. "Our students who would like to be in management need to be able to visualize data, pose questions about the data and interpret what's given to them," says McGaughey. That's just what Cal Poly is teaching them to do.

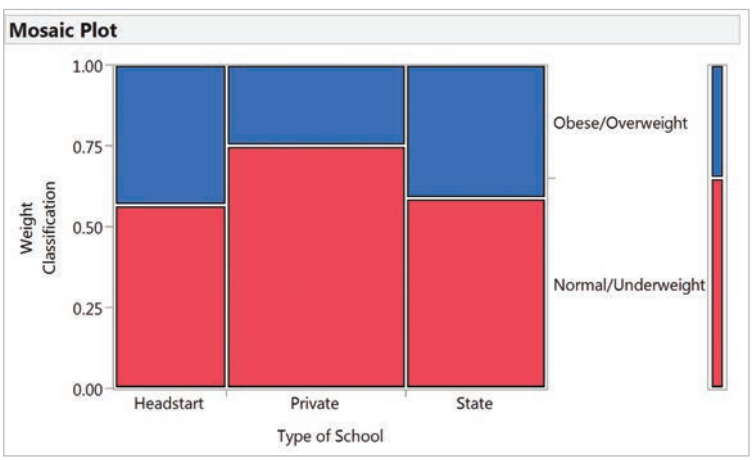
Cal Poly prepares its graduates to keep pace with the era of big data

Universities like Cal Poly now emphasize conceptual learning and simulation. Thirty years ago, this wasn't possible because not everyone had a personal computer. Today, every student can use applications like JMP to find the statistical significance of collected data.

In Cal Poly's statistics department, students work on realistic problems that emphasize the understanding of all aspects of statistics: the planning of sample surveys and designed experiments, the process of acquiring data, the careful analysis of information, and the communication of results and conclusions. McGaughey explains, "Now the data sets are large and interesting, whereas historically the data sets were really small. We're using real data to show students how decisions are made. We can give them a data set on cancer. We can appeal to engineers or psychologists with real data related to

STATISTICS IN THE REAL WORLD

"Students use this mosaic plot in a statistics lab in which they study the association between obesity in preschool-age children and the type of preschool they attend. The type of preschool is related to socioeconomic status," explains McGaughey.



Mosaic plot prepared in JMP. The data was collected as part of a research study on obesity conducted by Cal Poly Professor of Kinesiology Kris Jankovitz.

"This is an example of a real data set, which is both timely (study of obesity in children) and relevant (conducted by a researcher at Cal Poly), that students get to explore and think about in my statistics class. I use it in a lab setting where the students are presented with a research question: 'How does obesity relate to the type of preschool?' Then the students receive the data set and are asked to create a graph that answers this question, followed by a chi-square test of independence."

JMP, with its visual, drag-and-drop, no-coding-required paradigm makes it easy for students learning statistics to focus on analyzing and gaining insights from the data, not on learning algorithms.

their fields. And because of the computing power, we can show them the real applications of the data."

In 2014, the American Statistical Association (ASA) changed its guidelines for bachelor's degrees in statistics, updating those it disseminated in 2000. The new document suggests changes in curriculum and pedagogy designed to ensure that students entering the workforce or heading to graduate school have the appropriate capacity to "think with data" and to pose and answer statistical questions. Key points focus on the increasing importance of data science, real applications, more diverse models and approaches, and the ability to communicate effectively.

Changes in the classroom include:

- The addition of a minor in data science in fall 2015, a collaboration between computer science and statistics.
- Planning for more computing courses beyond existing classes in SAS and R. Classes will feature instruction that integrates these computing packages in a variety of statistical and communications fields.
- Giving students more opportunities to hone communication skills through oral presentations and discussions of statistical topics. The McKinsey research indicates the world needs more people who can talk to each other about statistics. Cal Poly agrees.

The road less traveled becomes the beaten path

The path to statistics that McGaughey and her faculty colleagues took was much less direct. Most came to statistics through a math, engineering or agriculture program. Some came through fields of epidemiology, physics or biology. A direct path simply wasn't available.

"I didn't know what I wanted to do. I started in engineering. I had a chemistry teacher in high school who inspired me. When I graduated with a bachelor's in chemistry, I wanted a master's in mathematics," McGaughey says. "Then a family friend shared what statistics could do. At that point, my fire was lit and I knew this was the direction I should go. The reason I came to Cal Poly is because its primary mission is undergrad education. My passion is teaching."

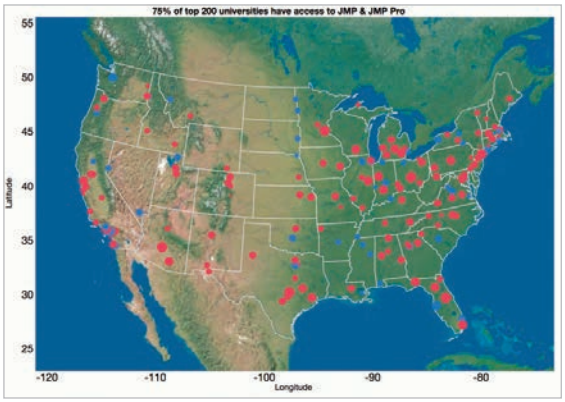
In recent decades, related topics like statistics, math and engineering were taught together. While land-grant institutions in particular usually had a statistics department, the demand for standalone statistics colleges wasn't there. "The mission of land-grant institutions was research in agriculture," McGaughey says. "That required a lot more resources than just one or two people designing their experiments and analyzing their data."

At Cal Poly, students used to find statistics courses through the math department, which offered courses in computer science, statistics, philosophy, biology, engineering, physics, astronomy and geoscience. The closest thing to a BS degree in statistics was a BS in math with a "statistics option." In 1969, the university created the Computer Science and Statistics Department and the Department of Statistics was ultimately founded in 1984.

Today the department's mission is to develop the next generation of statistics professionals while increasing the statistical literacy of Cal Poly students and the frequency of collaboration with their colleagues in related fields of study. "By having the students declare their majors upfront, we're building a solid foundation and sense of belonging that we hope will set them up to be successful," McGaughey explains.

600+ UNIVERSITIES LICENSE JMP®

More than 600 universities around the world license JMP and JMP Pro, including 75 percent of the top 200 universities in the United States. JMP has become an essential tool for teaching and learning statistics, and is currently used by more than 100,000 students in 1,000 courses across the US alone. This map shows the locations of US universities that currently have access to the software.



ONLINE

JMP partners with professors, students and researchers, offering a rich assortment of resources. Check it out: **jmp.com/academic**

Cal Poly has promoted learning by doing since 1903: **calpoly.edu**

Analytics cures a giant headache

JMP® is the 'active ingredient' for easing processing pains

What began 200 years ago as a storefront pharmacy in Edinburgh, Scotland, has grown into the world's leading supplier of opiate alkaloids, including morphine and codeine. Today, Macfarlan Smith Limited supplies pharmaceutical firms around the globe with the active ingredients for many best-selling pain relief medications.

But recurring production issues were causing a giant headache at a recently expanded facility. And no one was suffering more than David Payne, then the company's Head of Continuous Improvement.

Data analytics and data visualization – terms that weren't part of John Fletcher Macfarlan's lexicon in 1815 – ultimately helped Payne and consultant Andrew Ruddick identify and resolve the production problems. "Being able to resolve this has certainly been pleasing for a number of people," says Payne, who is now Head of Production for the Edinburgh facility.

A 'strategic growth area'

Growing demand had prompted Macfarlan Smith to increase the capacity at the Edinburgh facility where morphine is extracted from "straw," a term for the pelletized husks of opium poppies. The morphine is converted to codeine and sold to pharmaceutical companies.

The extraction facility operates around the clock. After the expansion, frequent processing issues were thwarting the increase in production that the renovation was designed to support.

"This is a key strategic growth area of the business," Payne explains. Failure to resolve the production problems "would

CHALLENGE

After an expansion to meet growing demand, repeated processing issues were thwarting the production increase that the renovation was designed to support.

SOLUTION

Tools for exploratory data analysis, modeling, profiling and what-if simulation did the detective work needed to root out and correct the issues.

RESULTS

JMP is now at work in other areas of the company.

have put in jeopardy a huge amount of work that had already been done."

Even the best efforts of a dedicated and experienced workforce couldn't ferret out the cause of issues that led to repeated process variation and manufacturing chokes.

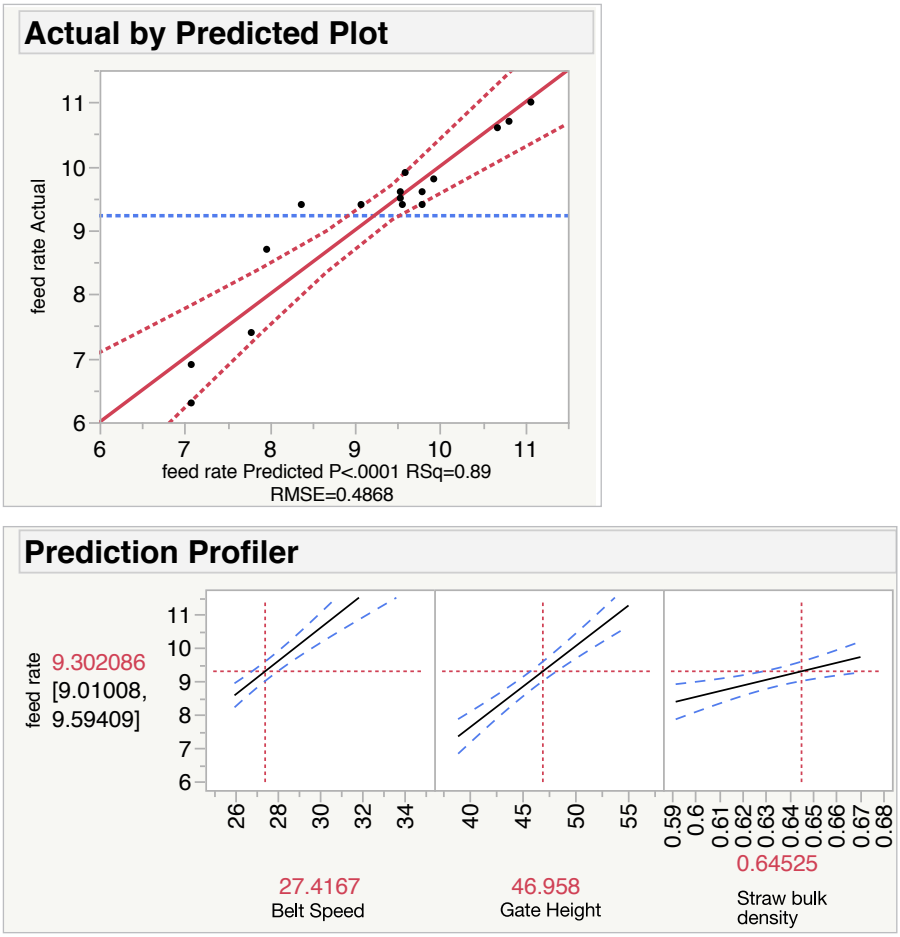
"Every day or two the plant would have to be shut down and started up again," Payne recalls. The effects were both financial and psychological. "When there are chokes or the equipment goes down, it can really begin to damage morale and cause frustration.

"Issues materialized in different places depending on how you were running the plant and the types of straw you were processing."

At the time, Payne was taking a class taught by Ruddick, co-founder of the UK consultancy Process Insight. The consultancy helps corporate clients improve operational efficiencies using Lean Six Sigma.

"Every trainee comes with a project for their company," Ruddick says. "In this case, David brought along this project because it was a critical problem for the business; he recognized it needed a fairly heavy dose of data analysis to figure out what was going on."

Process Insight uses JMP for data analysis, and Payne's project "was absolutely ripe for JMP," says Ruddick. The software is designed to do the detective work of sifting through vast volumes of data to reveal connections and trends that mere spreadsheets can hide. "There was plenty of data around, but it needed some careful exploration – discovery, really – to see the forest for the trees."



JMP models helped optimize the feed rate for poppy straw going into the extractor and helped illustrate the effect of changes to the speed of the belt carrying the straw and the height of the gate the straw passed through.

From flower to 'straw'

Deriving morphine from poppies is a multistep process. After harvesting, the seeds are removed from the flowers, and then the husks are milled and formed into pellets. The resulting straw is washed in a solvent to extract morphine. Then a purification process captures the morphine from the solvent. Another Macfarlan Smith facility then converts the morphine to codeine and other products.

The first goal of the project was to determine the optimal rate for feeding poppy straw into the extractor. When Payne and Ruddick used JMP to examine parameters they had assumed were responsible for the feed rate – weight changes in big siloes that stored the straw – the results were surprising.

"It became apparent very quickly that those weight changes weren't measuring what they thought they were measuring," says Ruddick. "They would never have known that without JMP. It was some of the modeling and visualization tools in JMP that got them to see that."

Payne and Ruddick used the JMP Profiler and other modeling features in the software to identify a more reliable set of variables, including the speed of the belt that feeds straw into the extractor and the height of the gate that controls the depth of straw on the belt.

JMP also helped adjust for the "almost infinite" variations among several types of poppies that fed the extractor, says Payne. "Because you're dealing with a natural product, that product varies,"

> "The success of this project, and specifically using JMP within it, elevated the importance of JMP at this site. More and more projects will be using JMP to do this sort of analysis."
>
> DAVID PAYNE

Payne explains. "And we didn't really have the understanding of how to set the plant up to deal with different types of straw."

Detective work

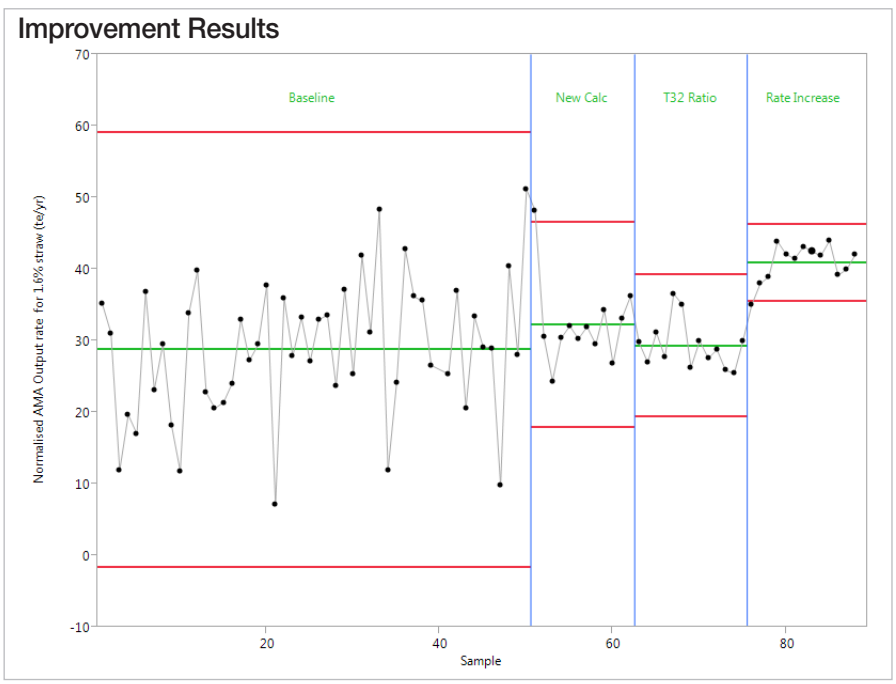
So Payne, Ruddick and the team built a JMP database that included thousands of bits of information – growing and harvesting practices, extraction procedures, equipment performance and more. "It was easy to bring the data together for analysis in JMP using simple scripting," said Ruddick. Then JMP modeling, profiling and what-if simulation tools did the detective work needed to determine a formula that optimized the delivery of poppy straw to the extractor.

Next, Payne and Ruddick focused on the next step in the extraction process: separating the morphine from the solvent that washes it out of the straw. Too concentrated, and there is an increased risk of choking the equipment, causing a shutdown. Too diluted, and the equipment is underutilized and yield slumps. Again, JMP helped refine the process to reduce the variability of flow and stabilize the morphine concentration.

Throughout the analysis process, JMP captured analysis results in interactive graphs and charts that let managers and operators alike see trends and patterns in the data and observe how changes to even a single variable affected the overall process. "It engaged the whole organization in problem solving by opening data analysis to everyone, not just statisticians," Ruddick explains.

"You can show them a profile and they understand that. It was important down into the organization and up into the organization to illustrate both what was going on and demonstrate how you could bring this back into control. It gave them a confidence that hadn't been there before."

Payne concurs. "We use JMP routinely now at that plant and will continue to do so going forward," he says. "The success of this project, and specifically using JMP within it, elevated the importance of JMP at this site. More and more projects will be using JMP to do this sort of analysis."



As process changes were implemented, JMP graphs helped measure incremental improvements in the extraction process.

ONLINE

How can you use statistical modeling to improve processes? Watch this Technically Speaking webcast: **jmp.com/improve**

Macfarlan Smith has manufactured and developed active pharmaceutical ingredients for 200 years: **macsmith.com**

Mining big data propels increased product miniaturization

Engineers apply JMP® and innovative thinking to create cutting-edge electronics

You may know Murata as a major supplier of high-end smartphone components – or you may recognize the company as an innovator in concept robotics. The MURATA BOY, a bicycle-riding robot who has delighted crowds at many a tradeshow, relies on state-of-the-art gyro-sensing and ultrasonic technologies to achieve seemingly impossible states of balance.

In fact, Murata views its concept robots as a way to showcase its many electronic manufactures. Among these, the highlights of Murata's catalogue include ceramic capacitors, inductors, gyro and ultrasonic sensors, filters, communication modules, ferrite beads and crystal units. In its industry, Murata is a global leader in product quality homogenization, miniaturization, high-performance, thinness and accelerated Kaizen cycles.

As increasing global innovation enlarges the market for leading-edge electronic components such as these, Murata acknowledges that manufacturing processes must keep up with the innovative spirit of the company's products. Unwieldy and inefficient processes limit growth.

So it falls to Murata's Monozukuri Enhancement Department to dig deep into the company's production systems data in search of clues as to how the company might design and implement strategic improvements. With massive

CHALLENGE

With advances in technology bolstering the demand for manufactured electronics worldwide, Murata sought a more creative, data-driven solution to optimizing not only its processes, but also its end products.

SOLUTION

Murata's engineers now rely on JMP to translate objective data analysis results into demonstrable on-site improvements.

RESULTS

With JMP, Murata streamlined elements affecting the ceramics calcination process shrinkage factor, creating an alternative control system by which to stabilize variation and reduce product size.

"By aggregating results in JMP, we've created an easy way to systematize operational evaluations across departments."

OSAMU SHIMOYAE

and complex data sets, however, data mining is no easy task.

Since adopting JMP in 2001, Murata has systematized its quest for improvements. By trolling the data for new ideas, the company keeps pace with its ever-evolving product offerings.

> "We continue to champion data mining as a means to achieve major process improvements. And that is how, ultimately, we continue to improve product quality."
>
> OSAMU SHIMOYAE

Murata mines production data for process improvement insights

There was no doubt by early 2001 that data mining would be the secret to Murata's continued success. With the inauguration of a devoted data analysis team the same year, Murata used SAS to construct a database integrating manufacturing data and information from across the entire global company. So while it is the role of SAS to accumulate and store data, it now falls to JMP to point at how that data might be used to resolve daily manufacturing challenges small and large.

At the helm of this effort is Osamu Shimoyae, Senior Manager of Production Innovation at Murata. "Data mining helps us to determine improvements through the analysis of mass production data generated during both the production process and site integration," says Shimoyae.

Murata's on-site data analysis team uses JMP to conduct multivariate analysis, making visible the relationships between variables by examining myriad possible variable combinations. Shimoyae says that, in rare cases (e.g., with highly novel products), unexpected relationships between factors emerge. However, many problems can be solved by quantitatively verifying the relationship between two variables based on a hypothesis within organized segments.

"We evaluate ROI strictly," says Shimoyae. "We use data analysis to prioritize challenges and tackle those for which we anticipate large improvement effects. By aggregating results in JMP, we've created an easy way to systematize operational evaluations across departments." Analysts then use site visits as a way of verifying trends they observe in the data. Sharing results on-site, analysts and engineers collectively brainstorm targeted, creative solutions.

Data analysis in real time

Murata believes that maximum improvement results when real-world science merges with data science – from production machinery sensor data to engineers' knowledge and know-how. And lucky for Murata, JMP has provided a useful forum for this fusion.

"The appeal of JMP," says Shimoyae, "is that the data is searchable." JMP can process data in real time, reflecting users' thought patterns and ideas. If you have a quandary, you can run it through JMP and have an answer straight away.

And such process improvements pay off, says Shimoyae. Noticeable results are produced on the factory floor, and these changes affect Murata's bottom line.

Analyses make increased product miniaturization possible

In addition to standard processes, new products account for around 40 percent of Murata's overall production, so measures to shorten product life cycles are essential. Accordingly, in order to offer products that meet market and customer needs, timely improvements are essential.

Variations in the shrinkage factor of ceramics arise due to issues in the calcination process such as laminate density, furnace temperature or friction between the calcinated product and substrate. As such, Murata collects the data generated in the production process and, using JMP, isolates variables that affect the shrinkage factor.

By understanding this relationship, Murata's engineers have succeeded in stabilizing variations in the shrinkage factor by improving the calcination process.

It was also possible to limit calcination-related shrinkage to the processing department's size requirements. "If there are large variations in shrinkage, increased gaps between parts are inevitable," says Shimoyae. "By stabilizing the shrinkage factor, we were able to narrow the gap between products and reduce product size as a result."

Standardization is also crucial. As are insights from Murata's on-site engineering teams. So Shimoyae and his colleagues use JMP as a means of standardizing methods for setting variables and narrowing down segments by systematically consolidating and documenting application methods.

"We continue to champion data mining as a means to achieve major process improvements," says Shimoyae. "And that is how, ultimately, we continue to improve product quality."

ONLINE

Murata is a global leader in advanced electronic materials: **murata.com**

See how others are using JMP in the manufacturing industry: **jmp.com/success**

Navigating road safety improvements

CHALLENGE

Visual stimuli both from inside and outside a vehicle are responsible for hundreds of distracted driving deaths each year. Improving the design of in-vehicle technologies and road signage to account for driver safety could save countless lives.

SOLUTION

Maryam Zahabi, an industrial engineer at North Carolina State University, uses JMP to analyze designed experiments that gauge the effects of on-road signage and in-vehicle technologies on driver attention allocation among both civilian drivers and police officers.

RESULTS

With the results of her experiments, Zahabi and her colleagues plan to provide the Department of Transportation with improved design and formatting guidelines for on-road signage. Her analysis will also help improve mobile computer terminal interface design for police cruisers nationwide.

An industrial engineer at North Carolina State University uses data to help the Department of Transportation reduce car crashes caused by driver distraction

According to research from the Centers for Disease Control, each day in the United States eight people are killed and 1,161 injured in motor vehicle crashes related to distracted driving. While cell phones are a common source of distraction among drivers, roadside signage can also draw drivers' attention away from the road. Directional guide signs as well as commercial advertisements are a considerable external distraction that can cause drivers to take their eyes off the road.

And driver inattention is not only a civilian problem; police officers are also prone to multitasking while driving. Of particular interest to Zahabi and her colleagues are the in-vehicle mobile computer terminals (MCTs) used by police officers and other emergency vehicle personnel. MCTs consist of a screen and keyboard that can be used to view and enter information, and to transmit data to peripheral devices like two-way radios. Officers are often called upon, through computer-assisted dispatching, to access and respond to on-screen information while driving, especially in emergency situations. But MCTs can also present a significant safety hazard for police officers who have to operate them while simultaneously operating a vehicle.

It is often the very tools that aim to improve the driving experience – and enhance the ability of police to enforce road safety – which present the biggest hazards. It's a conundrum – and a unique engineering challenge.

In order to design better technologies, engineers first need to understand how drivers respond – whether consciously or subconsciously – to a host of visual stimuli. By identifying those triggers that cause drivers to divert their concentration from the road, engineers can propose

[1] National Center for Statistics and Analysis, Distracted Driving: 2013 Data, in Traffic Safety Research Notes. DOT HS 812 132. April 2015, National Highway Traffic Safety Administration: Washington, D.C.

> "I didn't know that I could use a statistical tool to design experiments, but after taking a course on DOE, I learned how I could use JMP to produce different [experimental] designs. Now I use JMP for designing experiments as well as doing data analysis."
>
> MARYAM ZAHABI

new designs that decrease or even eliminate the greatest risk factors.

Human-systems engineering improves usability and safety by considering cognitive factors

Maryam Zahabi, a PhD candidate at North Carolina State University's Department of Industrial and Systems Engineering, has taken on that very challenge, using both quantitative and subjective data collected during experimentation to inform safety-related design improvements. The goal of the human factors area of research at NC State, Zahabi says, is to improve quality of life. "By designing on-road signs and in-vehicle technologies in a way that does not distract drivers from the task at hand, our research will improve safety for all people."

Zahabi came to NC State in 2013 from Iran, where she completed both her BS and MS at the prestigious Sharif University of Technology in Tehran. With a background in industrial engineering and ergonomics, Zahabi says she became interested in studying the cognitive human aspects of systems engineering after getting involved in research at NC State. "Coming to the department, I saw very exciting research projects in the human factors area – multitasking, driver distraction and usability in electronic medical records, for example." Ultimately, Zahabi decided to migrate her own research into the cognitive domain.

In the long run, research like Zahabi's aims to help people better navigate the world. "Ultimately," she says, "the findings of this research will be helpful in improving safety for both civilians and police officers."

Strategic design of experiments enables researchers to determine causal relationships in multifactor opportunity spaces

In the human factors field of engineering, Zahabi says, data analysis is the most important aspect of research. "We want to make sure that the findings of our research are powerful and can be generalizable. We need to make sure that our proposed enhancements could actually increase performance with statistically significant findings."

Industrial and systems engineers face complex multifactor opportunity spaces wherein a meticulous, methodical approach to experimentation is crucial to producing usable results. Information gathering in these research environments, however, can be challenging as relationships between inputs and outputs are difficult to spot. To ascertain such linkages, researchers like Zahabi and her colleagues manipulate the data in deliberate, strategic ways. To organize this process, many turn to a strategy often referred to as design of experiments (DOE).

DOE involves the adoption of a pre-specified design that enables researchers to model the ways in which factors independently or jointly affect a response. When she first arrived at NC State, Zahabi recalls, "I didn't know that I could use a statistical tool to design experiments; I always thought that software was only for data analysis. But after taking a course on DOE, I learned how I could use JMP to produce different designs…. And now I use JMP for designing experiments as well as doing data analysis."

Zahabi says JMP affords her the ability to create custom designs around specific questions, quickly and easily.

One software solution for many kinds of statistical analysis

In addition to DOE, Zahabi also uses JMP to run ANOVA, power analyses and correlations. Zahabi and her colleagues work with continuous (e.g., task completion time), binary (e.g., gender), nominal (e.g., physical exertion level) and ordinal (e.g., age range: young, middle, elderly) data, as well as subjective data from literature surveys and Likert scales. JMP, she says, "helps me to do different types of analysis – like parametric and nonparametric analyses and correlation, for example – very easily. And I also use JMP for data screening and the identification of outliers before applying inferential statistic methods."

ONLINE

Explore leading-edge engineering research at NC State's Edward P. Fitts Department of Industrial and Systems Engineering: **www.ise.ncsu.edu**

Read about how others are using designed experiments in engineering, science and manufacturing: **jmp.com/success**

Design of experiments brings order to the research cycle

Scientists in the pharmaceutical industry learn to consolidate steps, save money

Arlenda is a Belgium-based company of 24 statisticians and scientists who provide consulting in applied statistics and modeling for the life sciences industry, primarily for pharmaceutical clients.

Arlenda assists in every aspect of its clients' operations, from the discovery stage of a product through its manufacture. This involves the development of processes, assessment of stability, optimization and much more.

It's all very complicated, of course, and Arlenda's task is to simplify the process – to provide scientists with models that make sense and tools they can use to work independently with their data.

"Sometimes we write complex equations," says Bruno Boulanger, Arlenda's co-founder and Chief Scientific Officer. "But our clients don't want to see complex equations. They want to say, 'What's the consequence if I do this? What will my decision be based on this?'"

That, Boulanger says, is where "JMP comes into the loop." Using JMP, Boulanger and his colleagues communicate the outcomes of complex models in a manner that those with little or no background in statistics can understand. "They can visualize the outputs," Boulanger says. "And that's very important."

Screening designs streamlines processes, saves money

Boulanger has been a JMP user since the late 1990s, and one of his first initiatives upon launching Arlenda was to introduce his team to the benefits of using JMP for design of experiments (DOE). Today, he says, "We promote JMP whenever possible."

Arlenda's consultants now use a number of JMP tools regularly, and Boulanger calls the software's DOE and modeling capabilities fundamental to Arlenda's mission. Pharmaceutical companies are committed to investigating ways to conduct more reliable experiments, and Arlenda uses JMP to do that. It's all in

> "All the scientists and departments we introduce to JMP DOE and modeling continue to use it. They're convinced it works."
>
> BRUNO BOULANGER

keeping with what Arlenda refers to as a "lifecycle vision" of an experiment: Define your objective, determine your analytical approach, design your experiment, perform your experiment, gather your results, assert your conclusion.

"Research is a cycle," Boulanger says, "and you need to encapsulate everything." DOE brings order to that cycle, and JMP offers a full arsenal of tested classical DOE designs. The custom design tool allows users to answer specific questions without wasting resources. Then, once the data has been collected, it's easy to model, see patterns of response, identify active factors and optimize responses.

An area in which Arlenda regularly uses DOE tools in JMP is with vaccine manufacturers. The production of a vaccine understandably involves a number of complex processes requiring a great many materials. Each experiment is extremely expensive to conduct. Arlenda therefore often uses screening designs in the early stages of these experiments to identify those factors that have the greatest potential to effect desirable outcomes.

Designed for ease of use

"With design of experiments, you can better determine the most informative, and the cheapest, experiment to conduct," Boulanger says. "Scientists really like that."

What scientists also very much like is that JMP presents their data in a format that allows them to visualize its nuanced patterns and test relationships between variables themselves. "They don't want us to develop a model and then assure them that it will work," Boulanger explains. Researchers want to see the thinking behind the development of that process; they want to be knowledgeable. They can then take it from there. Simply put, JMP is easy to use. "It's been conceived from the user's point of view," Boulanger says.

Clients keep coming back

Boulanger contends that there is both an immediate and a long-term benefit in using JMP. The immediate benefit is that JMP allows you to perform fewer experiments to reach a conclusion, making it easier to predict how many experiments you'll need to perform. The long-term benefit is that it dramatically reduces the risk of failure in the future. "Because once you optimize a process using a design of experiments strategy," he says, "you then minimize the risk of producing a batch that's out of specification."

Although Boulanger isn't privy to information on how much his clients save as a result of using the service and support Arlenda provides, he's confident that those savings are substantial. "My metrics are that every company we've made our recommendations to has come back to us," he says. "And all the scientists and departments we introduce to JMP DOE and modeling continue to use it. They're convinced it works."

CHALLENGE

To help pharmaceutical companies establish more effective, less expensive development and manufacturing processes.

SOLUTION

In the course of its consultancy, Arlenda deploys a full suite of JMP features. Among the most frequently used, and enthusiastically received, are JMP design of experiments, modeling and graphic capabilities.

RESULTS

Arlenda's consultants provide scientists with an analytical toolkit based on JMP that can be adapted independently to future experiments and processes. As a result, Arlenda's clients keep coming back.

ONLINE

Arlenda's statistical experts serve clients in Europe, Asia and the US. Check it out: **arlenda.com**

Learn what researchers are doing with design of experiments in JMP: **jmp.com/doe**

Watch the JMP DOE video series with Princeton Professor Emeritus Stu Hunter: **jmp.com/doeplaylist**