JMP-START INTO TEACHING APPLIED STATISTICS

UKCOTS 2024, Manchester

Volker Kraft, JMP Global Academic Program

Abstract

Today's curricula often require students to learn a range of analytics skills, which are critical for all practitioners who want to learn from data. With the right software, learning these skills can be hands-on and engaging, allowing students to explore and analyze realistic data without struggling with a clunky or tedious statistics tool. JMP is interactive and powerful point-and-click software for solving real-world problems. It is ideal for engaging, hands-on teaching of relevant data skills in various fields and is also used by scientists and engineers at leading companies across the globe.

While the fundamental skills addressed in this session include understanding variation and uncertainty, we will also look at applications like data modeling, designing experiments and quality management – all from a student's perspective.

This interactive session will demonstrate how JMP can help to engage students' curiosity and teach relevant data skills which are most in-demand in industry today. We'll guide you through a series of brief demonstrations, so that you can directly experience the difference JMP can make for your course. Participants will receive a free license before the workshop, and the presenters will provide sample data and lead you through several hands-on examples in JMP.

We will also discuss best practices and share resources to support integration into modern statistics courses.

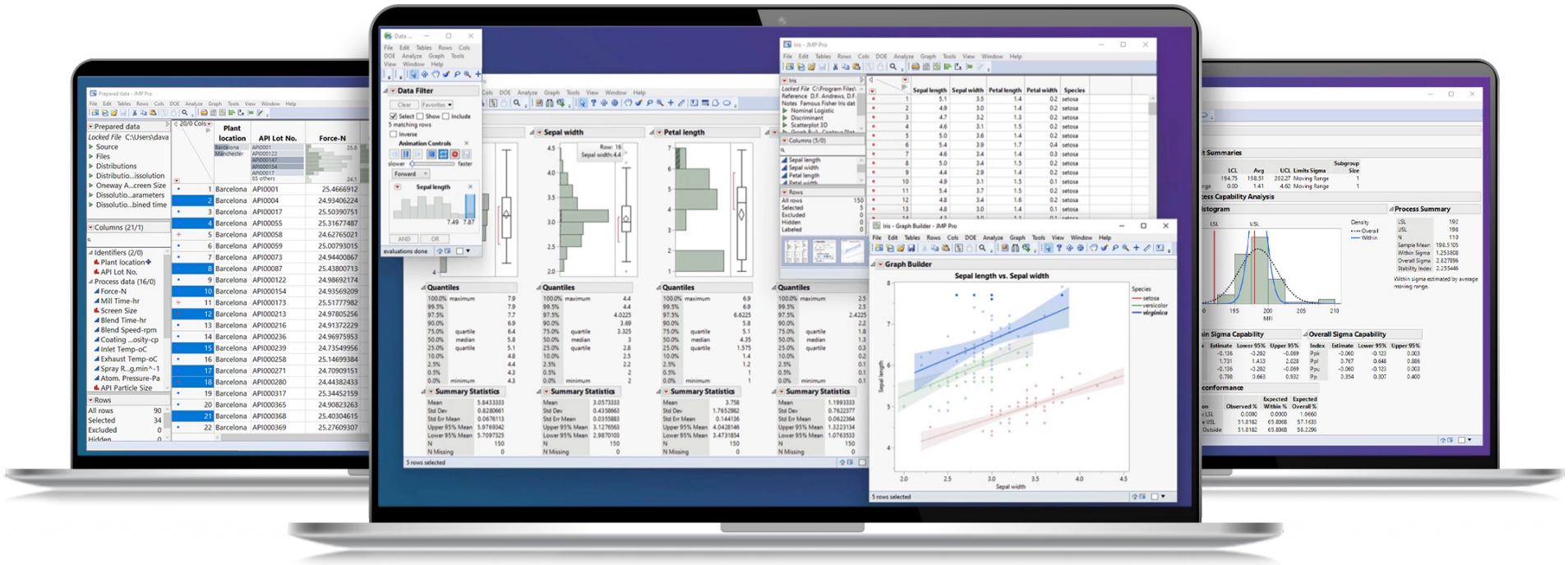
Agenda

1. Live demo: Teaching examples in JMP

2. Access for academic use:
   - Free resources
   - Free software

3. Discussion

1. Live demo:
Teaching examples in JMP



JMP® software combines interactive visualization with powerful statistics, without the need for coding.

Demonstration:

Enhanced Datasets

Datasets useful for demonstrating statistical techniques and JMP platforms. Complemented with descriptive storylines, exercises, and supplemental materials, these enhanced datasets are designed to engage students in the process of problem solving through statistical analyses.

› View Enhanced datasets

In addition, JMP comes with over 500 sample data sets included in the software, accessible through the JMP Help menu. Other datasets can be found in the JMP User Community and in JMP Public.

jmp.com/courses
> Introductory Statistics
> Enhanced Datasets

JMP® ENHANCED DATASETS

PENGUINS : SEX AND SPECIES COMPARISON

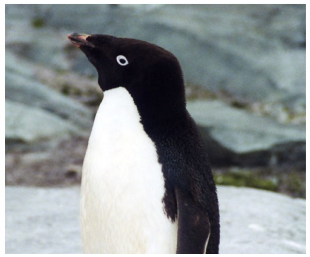
RELEVANT JMP PLATFORMS AND STATISTICAL TECHNIQUES

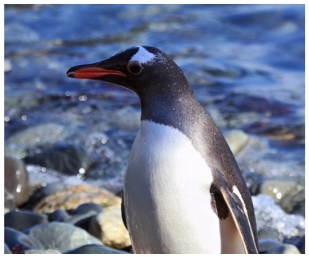
Graph Builder & Graphs :	Comparative Dotplots ; 3D Scatterplot
Multivariate :	Scatterplot Matrix
Fit Y by X :	One-Factor ANOVA
Distribution :	Univariate analyses, Confidence Intervals for the Population Mean, Tolerance Intervals for proportions of a Population.

PROBLEM STATEMENT

To better understand similarities and differences in morphological features between males and females of different penguin species, a researcher from Simon Fraser University in collaboration with The Palmer Station Antarctica Long Term Ecological Research Network obtained measurements on a sample of penguins from three different species [1].

The dataset consists of measurements across 4 physcial features of 34 male and 34 female Chinstrap penguins, 71 male and 72 female Adélie penguins, and 61 male and 58 female Gentoo penguins.





Demonstration:

Instead of watching the live demo at UKCOTS-2024, we suggest our on-demand resources:

- JMP in 2 Minutes (2 min, YouTube)

- JMP 101: Intro to JMP for Teachers (60 min)
- Intro to JMP for Students (60 min)
- New in JMP 18 and JMP Pro 18 for Academics (60 min, incl. Python integration and Deep Learning)

2. Access for academic use:
- Free resources
- Free licenses
- Free support

‘STIPS’ MOOC

Free, online, self-paced statistics course:

- Ideal for anyone wanting to learn fundamental skills around core statistical applications.
- Shaped by industry experts.
- Helps drive statistical adoption and comprehension.

www.jmp.com/statisticalthinking

Teaching Libraries

Teach with JMP

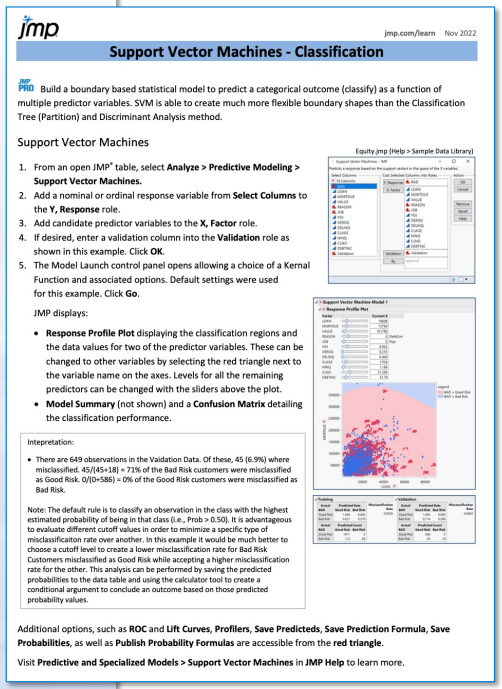
Use JMP's capabilities and teaching resources more effectively in your course

- JMP 101: Intro to JMP for Teachers
- JMP 101: Intro to JMP for Students
- Resources for Teaching with JMP
- Univariate Statistics and Probability
- Statistical Inference
- Exploratory Data Analysis
- Regression and ANOVA
- Advanced Regression
- Categorical Data Analysis
- Predictive Modeling
- Clustering
- Multivariate Methods, Pt. 1
- Multivariate Methods, Pt. 2
- Mixed Models
- Text Mining
- Statistical Quality Control
- Survival and Reliability
- Design of Experiments

JMP for Your Discipline

See how JMP's capabilities and resources align with your discipline

- Teaching Analytics in Chemistry and Chemical Engineering with JMP
- Teaching Engineering Statistics with JMP
- JMP Pro for Teaching Business Analytics and Data Science
- Better Teaching (and Using) Materials Science
- JMP for Teaching Statistics in Life Sciences



jmp

jmp.com/learn Nov 2022

Support Vector Machines - Classification

JMP PRO Build a boundary based statistical model to predict a categorical outcome (classify) as a function of multiple predictor variables. SVM is able to create much more flexible boundary shapes than the Classification Tree (Partition) and Discriminant Analysis method.

Support Vector Machines

Equity.jmp (Help > Sample Data Library)

1. From an open JMP® table, select **Analyze > Predictive Modeling > Support Vector Machines.**
2. Add a nominal or ordinal response variable from **Select Columns** to the **Y, Response** role.
3. Add candidate predictor variables to the **X, Factor** role.
4. If desired, enter a validation column into the **Validation** role as shown in this example. Click **OK**.
5. The Model Launch control panel opens allowing a choice of a Kernal Function and associated options. Default settings were used for this example. Click **Go**.

JMP displays:

- **Response Profile Plot** displaying the classification regions and the data values for two of the predictor variables. These can be changed to other variables by selecting the red triangle next to the variable name on the axes. Levels for all the remaining predictors can be changed with the sliders above the plot.
- **Model Summary** (not shown) and a **Confusion Matrix** detailing the classification performance.

Intepretation:

- There are 649 observations in the Validation Data. Of these, 45 (6.9%) where misclassified. 45/(45+18) = 71% of the Bad Risk customers were misclassified as Good Risk. 0/(0+586) = 0% of the Good Risk customers were misclassified as Bad Risk.

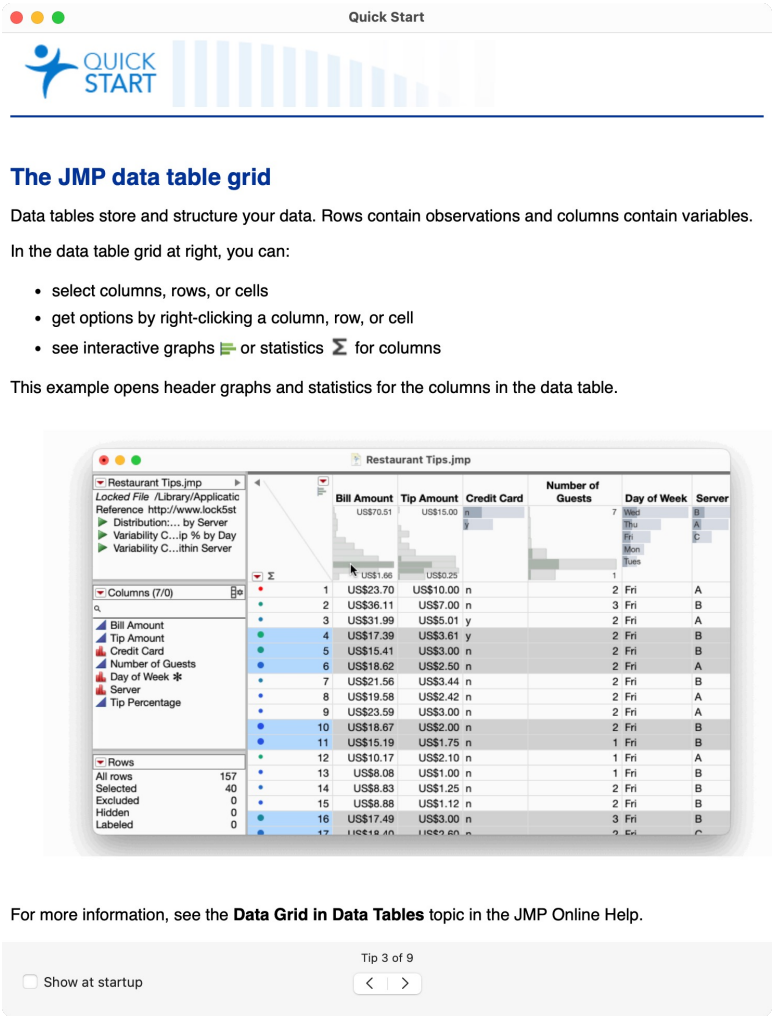
Note: The default rule is to classify an observation in the class with the highest estimated probability of being in that class (i.e., Prob > 0.50). It is advantageous to evaluate different cutoff values in order to minimize a specific type of misclassificaiton rate over another. In this example it would be much better to choose a cutoff level to create a lower misclassification rate for Bad Risk Customers misclassified as Good Risk while accepting a higher misclassification rate for the other. This analysis can be performed by saving the predicted probabilities to the data table and using the calculator tool to create a conditional argument to conclude an outcome based on those predicted probability values.

Additional options, such as **ROC** and **Lift Curves**, **Profilers**, **Save Predicteds**, **Save Prediction Formula**, **Save Probabilities**, as well as **Publish Probability Formulas** are accessible from the **red triangle**.

Visit **Predictive and Specialized Models > Support Vector Machines** in **JMP Help** to learn more.

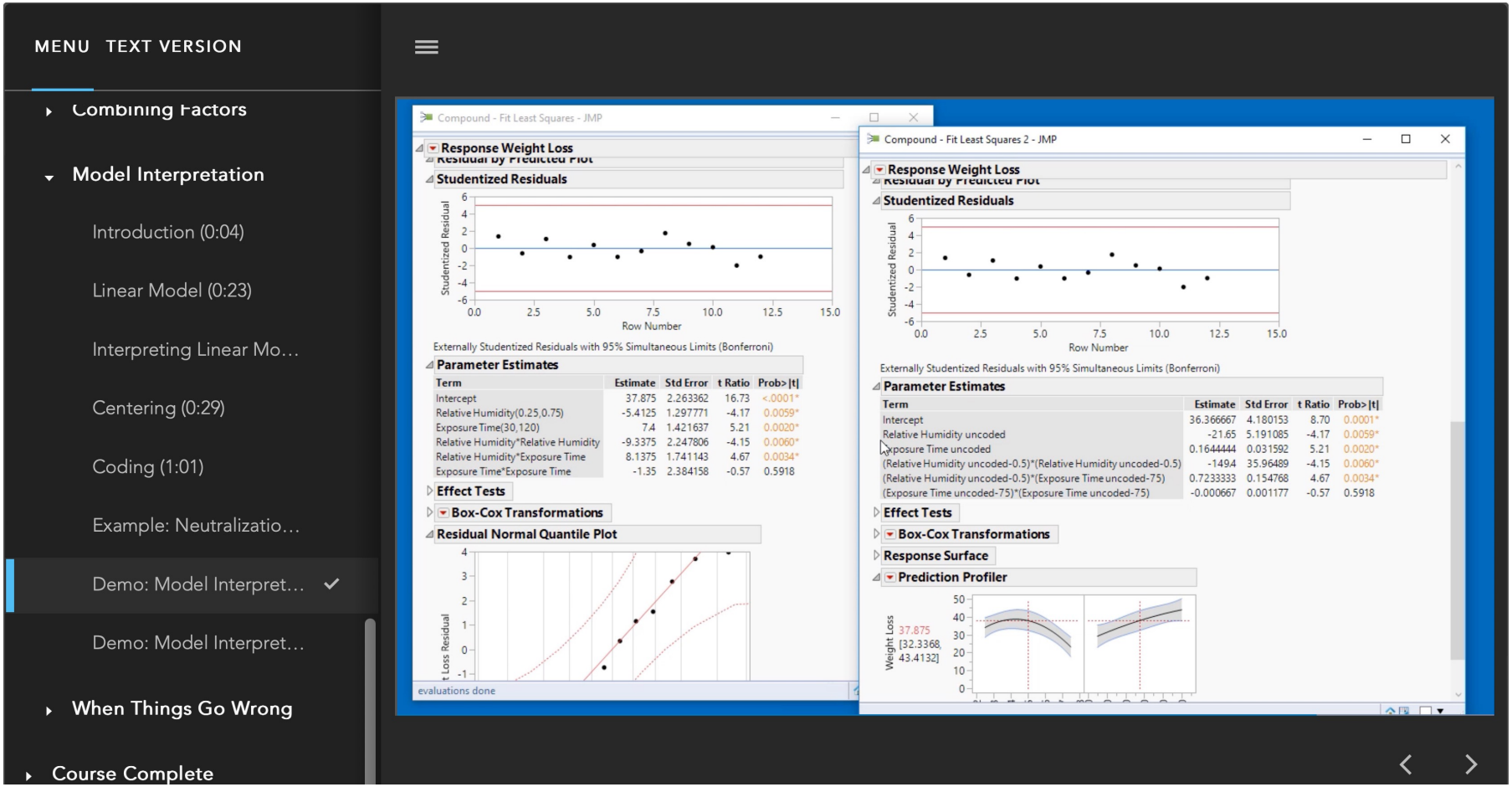
- Learning Library: 100+ "how-to" one-page guides and videos
  - *www.jmp.com/learn*
- Case Study Library: 60+ real-world problems with practical multistep solution paths
  - *www.jmp.com/cases*
- Academic Webinar Library: 40+ on-demand and live webinars
  - *www.jmp.com/webinar*
- Course Materials by course type
  - *www.jmp.com/courses*

Other Resources

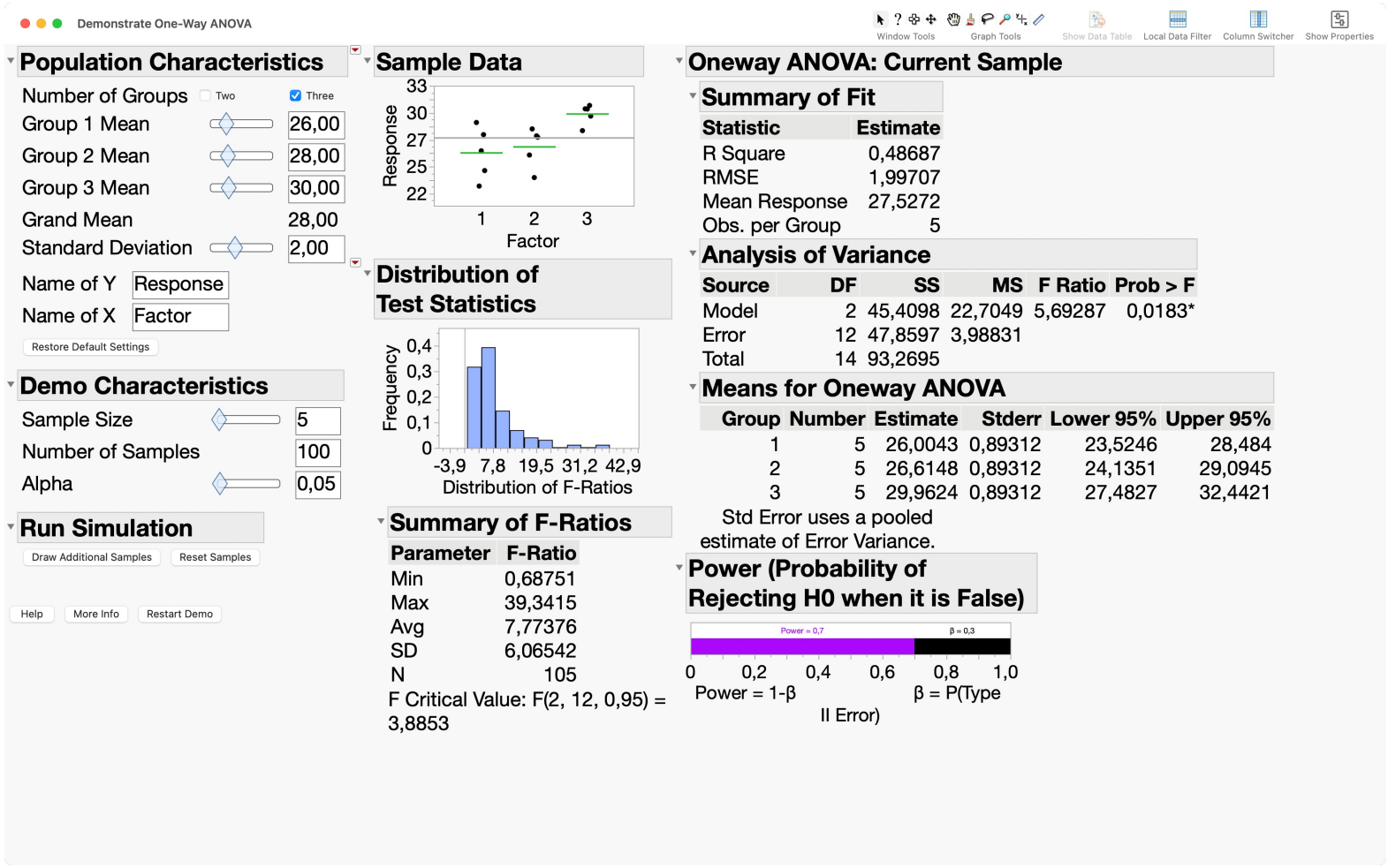


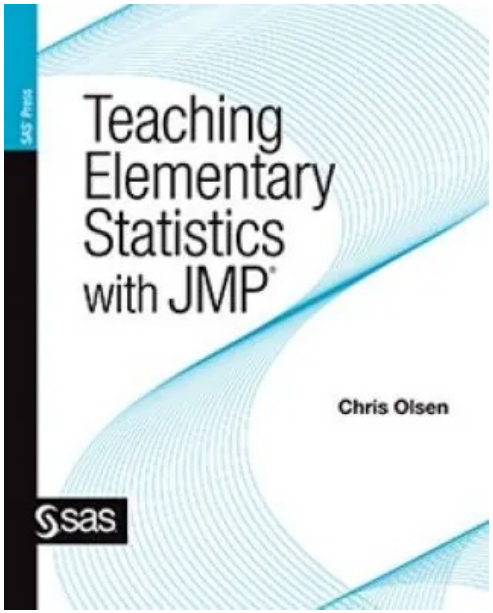
- Quick Start: Built-in tips to learn the basics of JMP
  - *Help menu*
- Teaching Modules: Interactive demos for statistical concepts
  - *Help menu > Sample Index*
- Enhanced Datasets: 20+ demos incl. background, exercises, links
  - *See course materials*
- On-demand courses: 7+ in-depth foundational online courses
  - *See course materials*

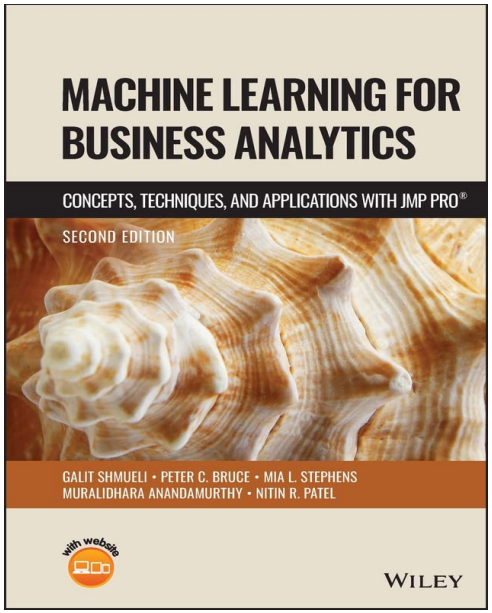
On-demand courses: Statistical Decisions Using ANOVA and Regression

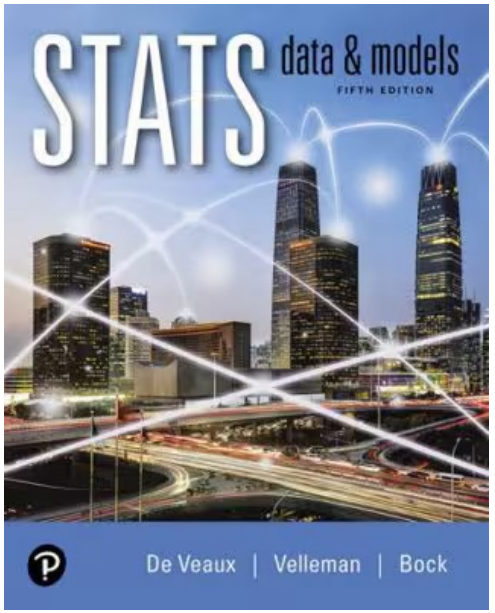


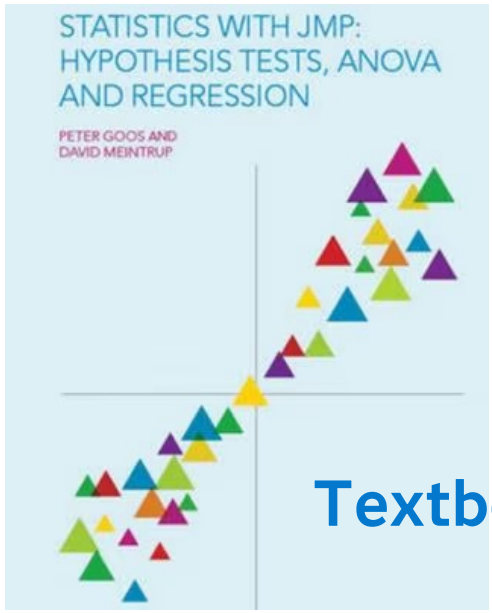
Teaching modules: One-Way ANOVA



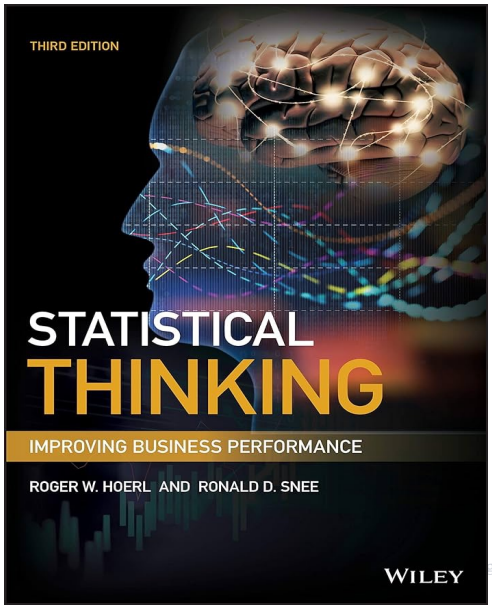


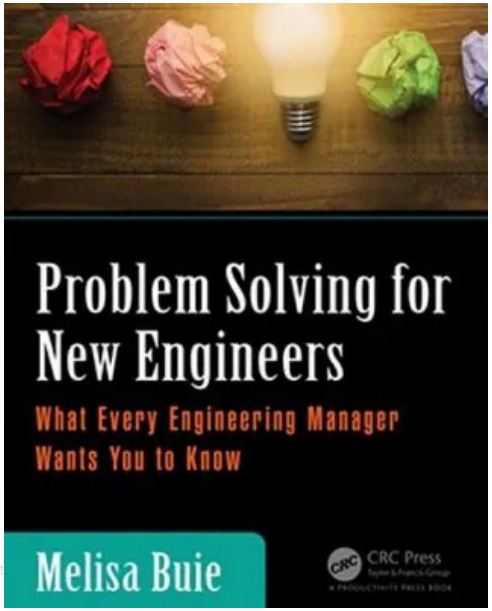


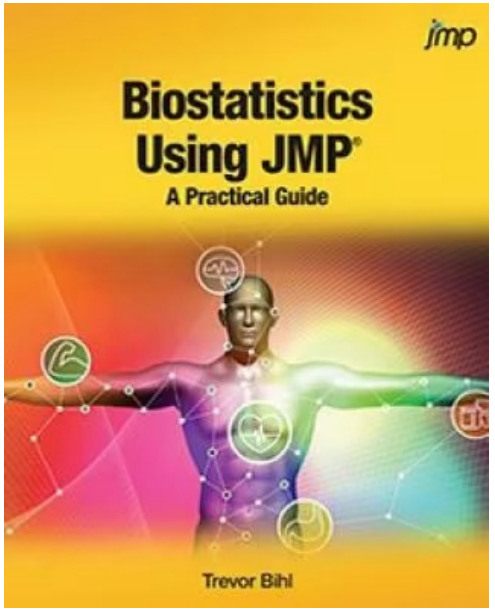


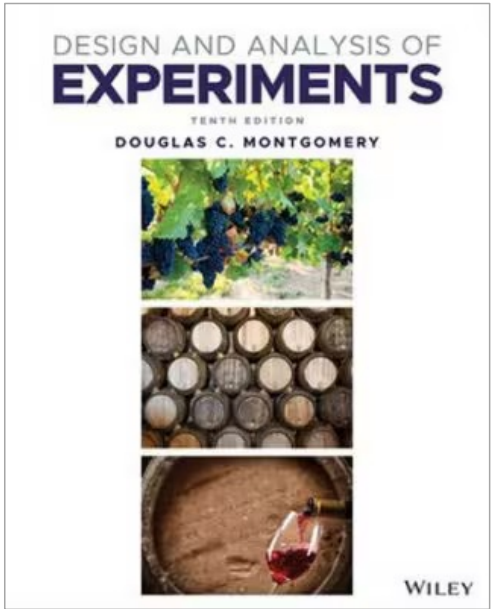


Textbooks



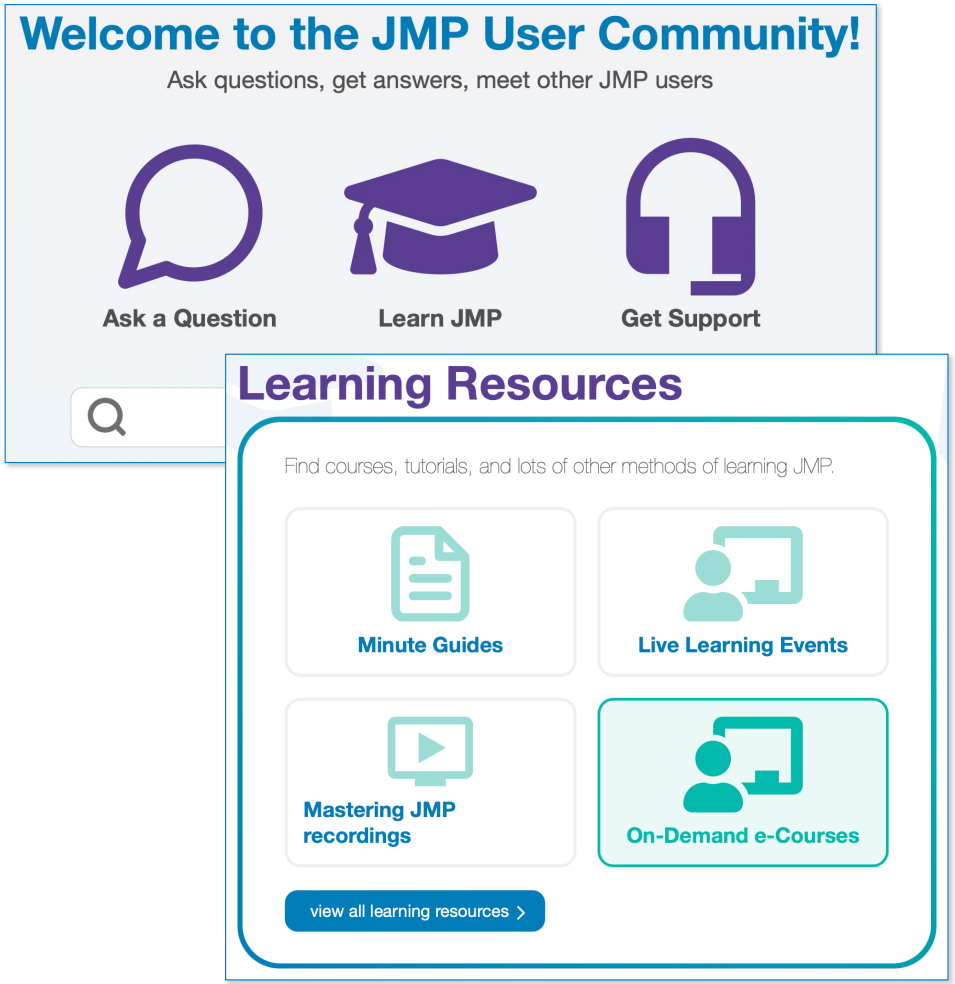






jmp

JMP User Community

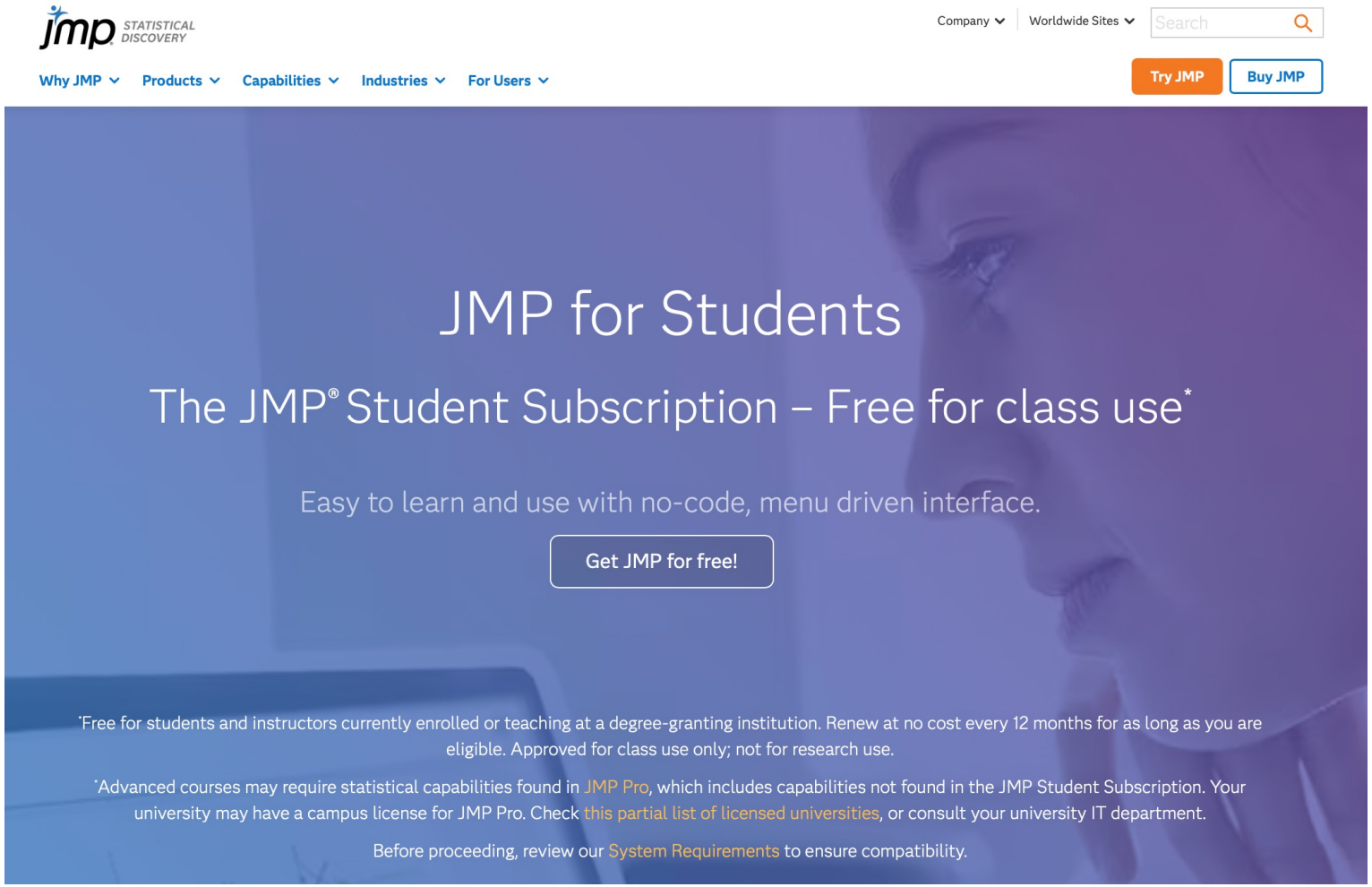


Networking with 50.000+ registered JMP users and experts:

- "Learn JMP" space
- Discussion forums
- Add-ins, data sets, etc.
- Blog posts
- Free in-depth e-courses

http://community.jmp.com/

Free access to JMP software: jmp.com/student



- Full version of JMP 18
- Free for class use
- Individual licenses for students and instructors
- 12 months, renewable

Other academic licensing options

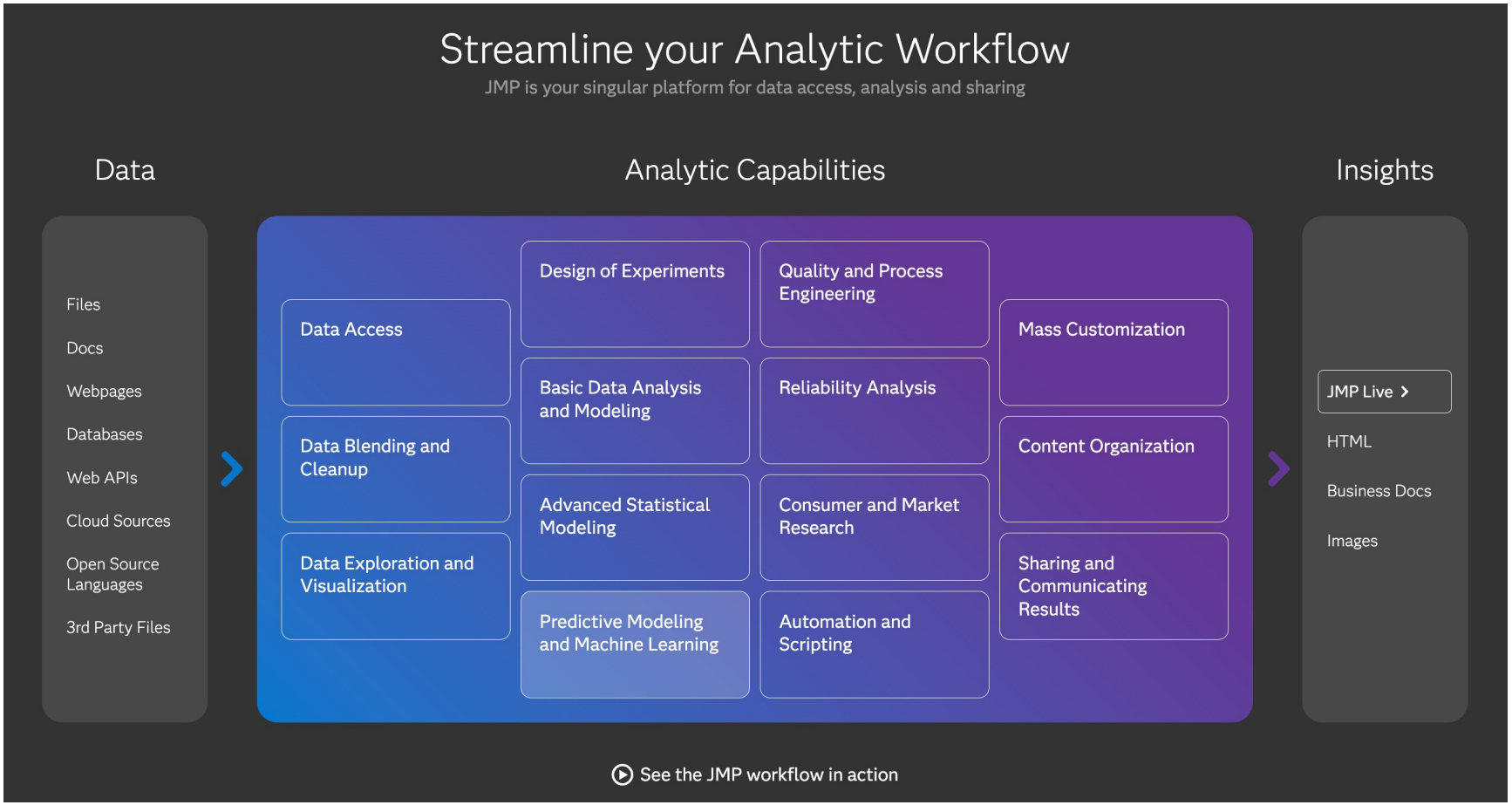
- Academic Suite License for teaching
  - Site licenses (campus-wide) at a degree-granting institution
  - Multi-user: Unlimited number of academic users, incl. home use; for classrooms, labs and research groups
  - JMP Pro* for Windows, Mac, and JMP for the Cloud; JMP Technical Support; free version upgrades; complimentary virtual and in-person workshops
    *) more ML, e.g. Deep Learning; Functional DA; Text Analytics; etc.
- Single-user JMP or JMP Pro license for research

Please send requests to academic@jmp.com

3. Discussion

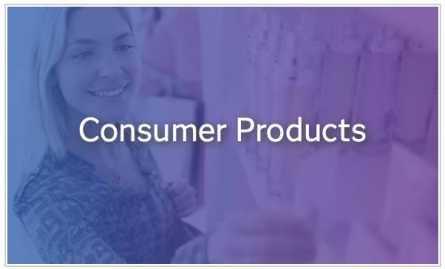
JMP provides all analytic capabilities…

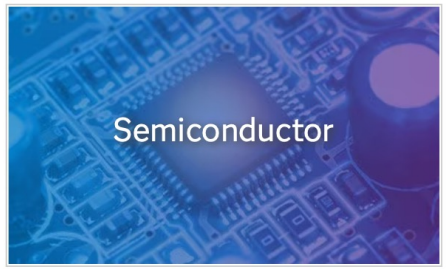
jmp.com/workflow



...solving real-world problems... (1/2)

jmp.com/en_us/industries.html





- Biotechnology
- Clean Energy and Conservation
- Government
- Industrial Manufacturing
- Medical Devices
- Medical Statistics

All 10 of the world's largest chemical companies use JMP.

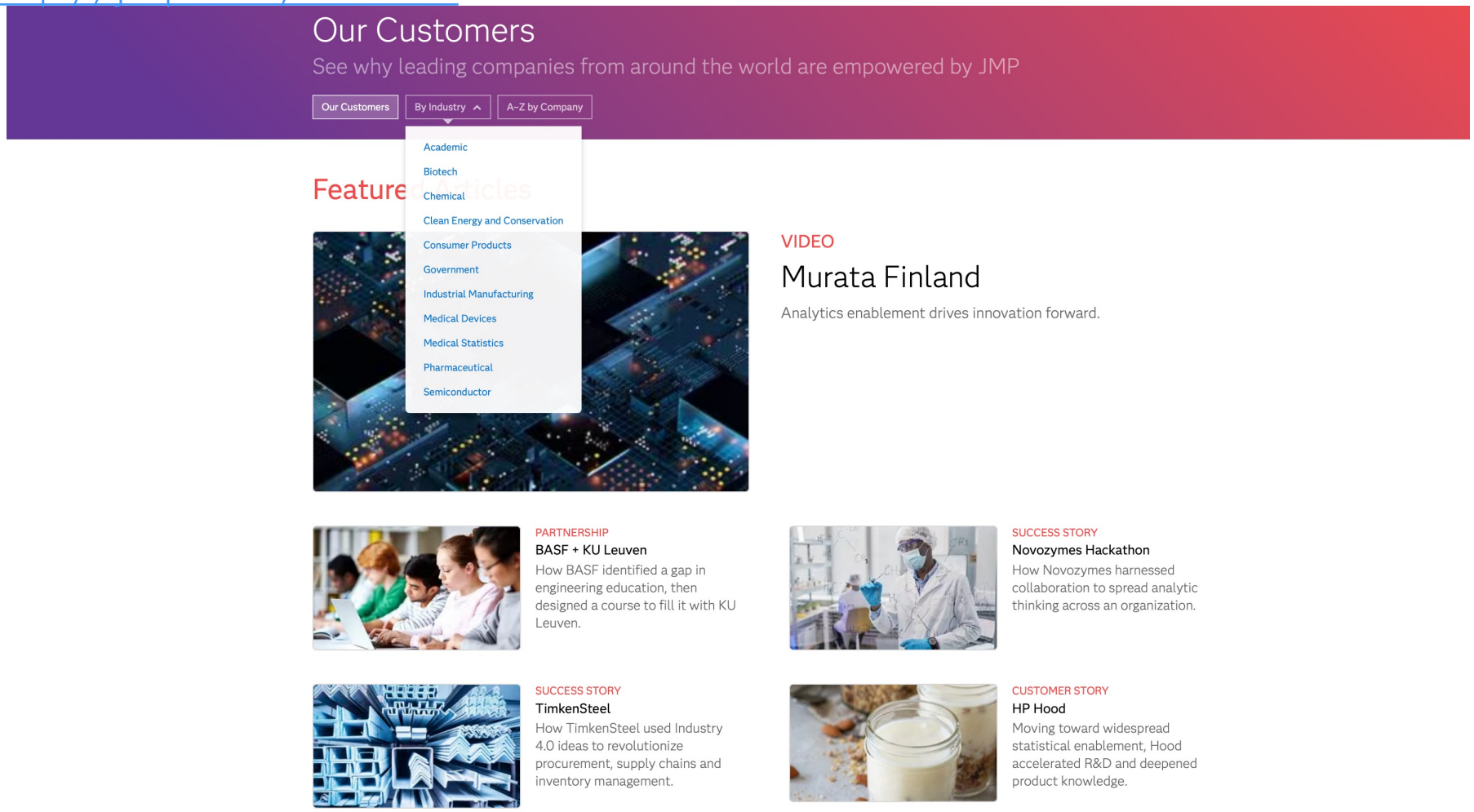
All 20 of the world's largest pharmaceutical companies use JMP.

All 25 of the world's largest semiconductor companies use JMP.

22 of the world's top 25 consumer product companies use JMP.

...solving real-world problems... (2/2)

http://jmp.com/success



… without coding

Benefits of a no-code solution:

- Time: No need for teaching the software or solving technical difficulties
- Mindset: Focus on statistical thinking rather than technical implementation
    - Data- and problem-driven rather than syntax- or command-driven
- Ease-of-use: Broader access to powerful analytics
    - Democratization of applied statistics

…but it is not code-less:

- Formula Editor & Full IDE for scripting in JSL = JMP Scripting Language
    - All analyses can be saved in JSL (created by JMP)
- Model scoring code can be exported in C, Python, SQL, SAS, JavaScript
- Optional integration using Python*, R, MATLAB, SAS, etc. (*embedded)

GAISE recommendations and JMP

Guidelines for Assessment and Instruction in Statistics Education (American Statistical Association 2005), including recommendations for introductory statistics courses on college and pre-K-12 levels in the US. Updated in 2016.

1. Teach statistical thinking
   - Teach statistics as an investigative process of problem-solving and decision-making
   - Give students experience with multivariable thinking
2. Focus on conceptual understanding
3. Integrate real data with a context and purpose
4. Foster active learning
5. Use technology to explore concepts and analyze data
6. Use assessments to improve and evaluate student learning

*) GAISE College Report ASA Revision Committee (2016): *"Guidelines for Assessment and Instruction in Statistics Education College Report 2016"* available at http://www.amstat.org/education/gaise.

Thank you!

volker.kraft@jmp.com

Please get in touch for a personal demo or discussion.

Recommended White Paper

The Integration of Big Data Analytics into a More Holistic Approach

By Roger W. Hoerl

The advent of new analytical methods like neural networks and bootstrapped forests has unlocked a range of exciting new possibilities when it comes to large data sets. Implementing these advanced methods, however, is not always as easy as it would seem. And that's where JMP can help.

In this whitepaper, statistician Roger W. Hoerl argues that JMP and JMP Pro provide workflows that help users avoid big data mishaps by getting back to the fundamentals of data quality. Beginning with an outline of several more commonly used methods, Hoerl then makes a case for integrating data quality checks in the analytical process. To assist with this synthesis, he identifies the core elements of a data pedigree and outlines a step-by-step holistic approach that applies traditional quality concepts to modern big data applications.