



Integrating SAS JMP® into the Research Process Moves Analytics from the Statistician's Desktop to the Laboratory Benchtop.

John P. Davies Jr., Dr. Teri Lalain-, Edgewood Chemical Biological Center, APG MD 21010

RESEARCH PROCESS FLOWCHART

Abstract

The Decontamination Sciences Branch (DSB) of the U.S. Army Edgewood Chemical Biological Center (ECBC) makes an effort to spread statistical knowledge and awareness throughout the organization. Experience shows that it is especially beneficial for those who are running experiments and generating data to be proficient in basic statistical tests such as t-test, ANOVA, and Tukey-Kramer. The DSB uses SAS JMP® and has incorporated it into all levels of decontamination research ranging from ANOVA to multivariate regression analysis and DOE. In addition to the analysis of single day tests, JMP® was recently utilized for a methodology validation using ISO 5725. The interactive capabilities and dynamic visual analytics were useful not only as a tool for statistical analysis but also as a vehicle to help permeate statistical understanding throughout the DSB from principal investigators to laboratory technicians. Because JMP® has a wide range of capabilities, from advanced analysis to univariate distribution functions combined with embedded help features, it supports the advanced or novice analyst. This flexibility gives JMP® a wide appeal that makes it particularly well suited for use throughout all levels of the organization. The ability to rapidly view data from many different perspectives, make discoveries, and quantitatively support conclusions is both satisfying and empowering to the individual. As more individuals experience hands-on data exploration, data analysis evolves from qualitative generalizations ("gut feelings") toward analysis based on numerical statistical calculations.

Introduction

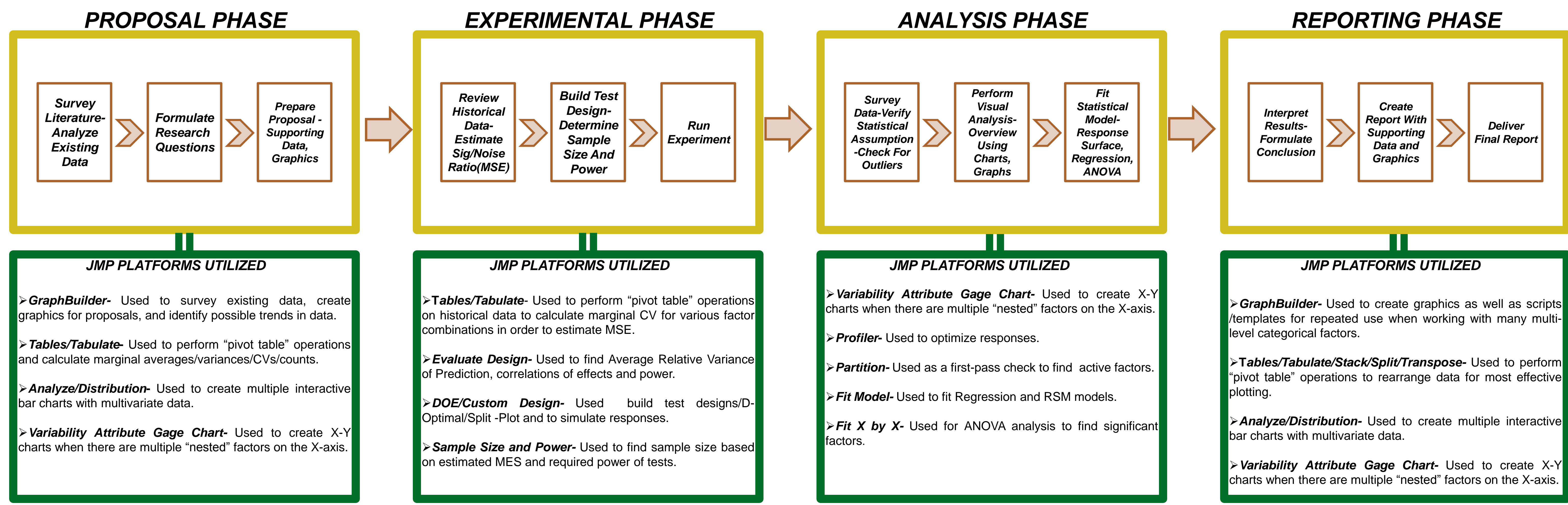
The DSB of ECBC performs chemical warfare agent decontamination testing over a wide range of conditions. Experimentation is of a highly multivariate nature and may involve numerous factors relating to substrate materials, temperatures, contamination times, decontaminant formulations, and decontaminant/agent concentrations. With each of these factor combinations there may be several measured responses of interest. Additionally, covariate values in the form of atmospheric fluctuations and other uncontrollable factors are also recorded. The multitude of factors and responses creates a challenging multivariate environment where efficient data management and analysis become essential. The means of distributing the raw data and analysis within the organization becomes a critical aspect of the research process flow. ECBC utilizes the SAS JMP statistical software package to meet these challenges. JMP software has been integrated directly into the research process flow from start to finish and is utilized across many levels of the organization by program managers, chemists, experimental designers, laboratory technical personnel, and statistical analysts.

Benefits of Integrating Analytics into the Research Process

- Having a standard analytical software eliminates the need for multiple imports/exports in and out of various spreadsheets or databases and allows for raw data, analyses, and ideas to be quickly shared. This is done in JMP through the use of shared data tables with scripted analysis attached directly to the data table. This allows for the thought process to flow between departments as one analysis builds on top of the other.
- Graphics and analysis techniques become standardized through the use of JMP scripted analysis and graphic templates. This increases efficiency and reduces the time from laboratory experimentation to final report.
- Providing laboratory personnel with an analytical tool allows them to perform real-time preliminary analysis of the data which can be particularly beneficial since they have the best "feel" for the non-tangible aspects of the data.
- Having an integrated statistics package in which people are proficient with the basic analysis platforms creates a foundation for teaching the more complex forms of statistical analysis.
- Providing all levels of the organization with "hands-on" data analysis increases the awareness of the importance of statistical concepts such as randomization, blocking, balanced design, and orthogonality.
- Integrating statistical software into all levels of the organization fosters an environment of quantitative thought where decisions are more likely to be data based.

Summary

There are numerous benefits of having a single statistical software package integrated throughout the research process flow. Perhaps the most valuable is that all levels of the research organization are given the ability to share data and statistical analysis in a highly efficient manner. In this way JMP effectively becomes a "conduit" for the transfer of data, ideas, and discoveries between personnel from all phases of the research process. When individuals from all levels are engaged in analytics the organization as a whole becomes more focused on the goal of transforming raw data into useful information.



JMP FACILITATES PROGRESSION OF DATA ANALYSIS

Research Question: What are the effects of test parameters /conditions on the decontamination performance?

Example #1

Principal chemist utilizes the JMP interactive **Analyze Distribution** function to analyze historical data and identify potential factors of interest. A JMP table serves as a benchmarking test plan and is passed to the analytical lab for execution.

Laboratory personnel run the experiment and Use JMP **GraphBuilder** to make an on-the-spot initial analysis and visually identify a relationship between the "A" level of the Decon factor and Temperature. The **GraphBuilder** graph is attached to the original data table (scripted) and passed on for a more detailed statistical analysis.

The Statistical Analyst runs the JMP **Fit-Model** platform and creates leverage plots to confirm that Temperature is statistically significant at only the "A" level of Decon.

The principal chemist receives the JMP data table with the various model fits attached as a scripted file. The **GraphBuilder** platform is used to create graphics that can be inserted directly into the final report.

Research Question: What is the most efficient experimental design to identify active factors and then optimize decontamination?

Example #2

Analytical lab executes the experiment and performs an initial analysis using the JMP platform **GraphBuilder** for visual overview, **Table/Tabulate** for marginal averages and **Fit X-Y** ANOVA for ordered differences reports. This initial analysis in the lab provides the opportunity for follow-up runs or investigation into outliers. The collected data and initial analysis are attached to JMP tables and passed to the statistical analyst.

The statistical analyst reviews the initial assessment from the lab and then runs the JMP **Partition** platform to quickly identify influential factors. The modeling platforms such as **Fit Model** and **Multivariate** analysis are then used to fit a model to the data. The **Profiler** platform may then be used to determine the factor settings that optimize the response of interest. The original JMP tables with the analysis and graphics attached in the form of scripts are passed on to be integrated into the final reports.

The principal chemist receives the JMP table and analysis and selects graphics to be inserted directly into the final report. Graphics are automatically regenerated from previously scripted analyses. The JMP **Variability Gage Plot** is often found to be useful for showing multivariate results.

Research Question: From a Quality Control standpoint, what are the root-causes of experimental noise?

Example #3

Quality control reviews the JMP tables/analyses of the positive control performance. The performance of the positive controls associated with particular combination of test factors is flagged as having an apparent non-random pattern in the data.

Data analyst runs JMP **Control Chart Builder** platform to generate control charts of the positive control response at the selected factor combinations that were identified by QC.

Experimental designer evaluates control charts and uses JMP DOE **Custom Designer** platform to design an experiment to identify which factors are impacting the response of the positive controls. The Design **Evaluation** platform feature **Color Map of Coefficient Correlations** and **Prediction Variance Plot** among others are used to assess the capability of the DOE design. The completed design is passed on to the laboratory for execution.