

# **Life Data Analysis with JMP® 8**

SAS Discovery Conference 2009

August 2009

**Edward Kram, P.E.**

President

Blue Arc Energy Solutions, Inc.



# CONTENTS

---

<b>1 CHAPTER 1 STRATEGIC OBJECTIVES .....</b>	<b>1-1</b>
Mission Clarity .....	1-1
Review of Raw Data Quality.....	1-1
Assessing Minimum Data Elements .....	1-2
Understanding Life Data Characteristics .....	1-2
Censoring .....	1-4
Truncation.....	1-6
Check Assumption Validity with Subject Knowledge.....	1-8
Repairable vs. Non-Repairable Systems.....	1-8
Distinguishing 1 <sup>st</sup> Failures of Repairable Systems.....	1-8
Censoring Must Be Non-Informative .....	1-8
Independent Failure Modes.....	1-9
Underlying Failure-Time Distribution Known vs. Unknown.....	1-9
Selection of Appropriate Statistical Tools.....	1-9
Chapter Summary .....	1-9
 <b>2 CHAPTER 2 TACTICAL OBJECTIVES: JMP® SEQUENCE .....</b>	<b>2-1</b>
Distribution Analysis Platform.....	2-1
ANOVA Analysis Platform .....	2-3
Graph Builder Analysis Platform .....	2-5
Tabulate Analysis Platform.....	2-6
Survival Analysis Platform.....	2-7
Competing Cause Analysis .....	2-8
Life Distribution Platform .....	2-9
Fit Life by X .....	2-10
Recurrence Analysis .....	2-11
Survival and Reliability Analysis II.....	2-12
Chapter Summary .....	2-12

<b>3 CHAPTER 3 PRESENTATION OF RESULTS.....</b>	<b>3-1</b>
Using JMP® Features to Save Results .....	3-1
Save to Table Scripts .....	3-1
Journals .....	3-2
Tune the Output to the Audience Preference When Possible.....	3-4
Visual Output .....	3-4
Statistical Theory within Text Results .....	3-5
<b>4 CHAPTER SUMMARY &amp; CONCLUSIONS.....</b>	<b>4-1</b>
Summary .....	4-1
Conclusions.....	4-1
Recommendations for Future JMP Features .....	4-1
<b>A APPENDIX: BIBLIOGRAPHY .....</b>	<b>A-1</b>

# LIST OF FIGURES

---

Figure 1-1 Complete Data.....	1-3
Figure 1-2 Right Censored Data .....	1-4
Figure 1-3 Interval Censored Data.....	1-5
Figure 1-4 Left Censored Data .....	1-6
Figure 1-5 Truncated Data.....	1-7
Figure 2-1 Distribution Analysis .....	2-2
Figure 2-2 Distribution Analysis for Initial Data Quality.....	2-3
Figure 2-3 ANOVA Example without Significant Mean Difference.....	2-4
Figure 2-4 ANOVA Example with Significant Mean Difference.....	2-5
Figure 2-5 Graph Builder .....	2-6
Figure 2-6 Tabulate.....	2-7
Figure 2-7 Survival Analysis .....	2-8
Figure 2-8 Competing Cause Analysis .....	2-9
Figure 2-9 Life Distribution.....	2-10
Figure 2-10 Fit Life by X.....	2-11
Figure 3-1 Save Results to Table .....	3-1
Figure 3-2 Naming Saved Results in Table .....	3-2
Figure 3-3 Save Results to Journal Example.....	3-3
Figure 3-4 Visual Output: MLE Fit Example.....	3-4
Figure 3-5 Statistical Theory Presentation Example.....	3-5



# 1

## CHAPTER 1 STRATEGIC OBJECTIVES

---

Reliability analysis is used in applications such as basic product reliability research, warranty and product liability, biological research, insurance, maintenance strategy, manufacturing design, and quality improvement. As a result a wide variety of statistical tools are available within these methods to practitioners.

The overall problem approach is described in this chapter includes some key elements to using JMP 8 successfully for reliability analysis applications. JMP 8 tools are very powerful for a wide range of uses. The intent of the chapter is to provide a general approach that works for a variety of end use applications.

### **Mission Clarity**

Understand what the end objective of the analysis is. This sounds simple but it can be lost in the rush to problem solve. To gain clarity it is necessary to get specific, for example: Is the mission's objective to define reliability over the entire lifespan of the subject? Or is the mission to focus on a specific life period, e.g., late life, or on a specific failure mode (cause)? To what level of certainty does the mission need to quantify the behavior? Will the application of results include a prediction of future behavior or failure quantiles? What is the risk with error and imprecision, e.g., is someone's life at risk, or is the risk purely economic, etc.?

Committing the mission objectives writing is a good practice and provides a tool to evaluate mission success upon work completion. Mission clarity also provides the ability to assess the raw data format and quality for suitability to the task. Other benefits of mission clarity include the overall efficiency of the tactical analysis sequence and goal orientation for the presentation of results.

### **Review of Raw Data Quality**

Get a good look at the data structure and quality to assess whether it is sufficient to conduct the analysis and achieve the objectives. Since reliability analysis quantifies the failure-time distribution, the most important data quality feature is its ability to quantify time-to-failure observations in the data. It is a good practice to review the raw timing data for completeness, appropriate format, and precision with respect to the mission objectives.

How many of the data are missing date information? Is the time to event data directly available in the raw data or does it require calculating from a start and end time/date? Is the timing data field of consistent format? Clearing up time/date formatting issues prior to importing the data into JMP precludes omission of data in the raw data import conversion and is a good practice to avoid backtracking. Or perhaps the raw data is in only discernable to monthly resolution but finer increments are needed for the mission. Are consistent timing formats used for both surviving units and failures?

An initial strategic look at the raw data quality will better position the analyst to evaluate the ability to complete the mission with the raw materials (data) given. Perhaps the mission requires 5% certainty but the data lacks date information for 25% of the observations. Perhaps the mission needs to predict failures within 10% each day, but the time/date resolution is at best monthly. These types of reviews are prudent at the front end of the mission.

### ***Assessing Minimum Data Elements***

There are some minimum data elements that should be present in the raw data. As discussed in the last text section, timing data is a minimum element. This time to failure data can take the format of an elapsed time field. It can also take the form of start and end time/date fields. This generally requires conversion to a single time to event field, but in the case of interval data, JMP can analyze with a start and end date data (assuming consistent format).

Unique subject numbering, e.g., unit serial numbers, are typically needed especially if the raw data sources are separate (failure history originates in one source and in-service (surviving) unit history originates from another data source) to preclude erroneous sample sizing or inaccurate assignment of the failure records to unit histories. This is a minimum requirement when the subjects are repairable. This enables analysis to establish sequential service segment history, i.e., down time (perhaps repair time also), and unit traceability when subject units are portable relative to the service position.

Cause of failure information, i.e., failure mode, is required for some reliability analysis platforms. Although it may not be a characteristic of the data for some applications, e.g., warranty data may contain the failure observation without cause of failure retention. Cause information should be reviewed for text validation, since collection may be limited in scope and subject to invalidated entry, e.g., misspelling, acronyms, inconsistent terminology, etc. Non-validated text entries may require cleanup (See JMP **Recode** features {underneath **Column**}) to preclude division of failures based upon trivial text differences.

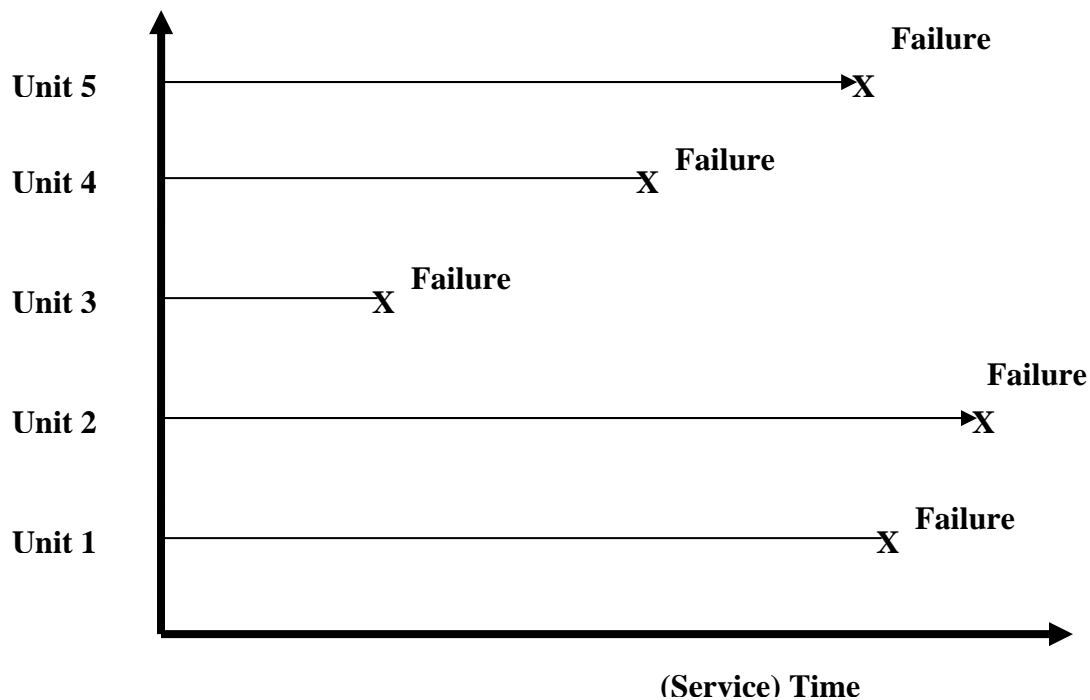
### ***Understanding Life Data Characteristics***

Unlike data from an enumerative study, in which the population is fixed or a representative sample from the population is assessed to draw inferences about the centrality of the population, life analysis is an analytical study in which the subjects for observation are generated over time. The inferences from analytic studies such as this assume that the process is constant. The process could be a manufacturing process, or on the flip side the process could be the service exposure.



There are two basic statistical classifications of life data types. Neither of these classifications imparts any assessment of data quality, which is a separate characteristic discussed briefly in prior text. Rather, “complete data” describes life data where all of the observations have exact failure times. “Incomplete data” describes life data in which some of the observations do not have exact failure times.

“Complete data” examples include a destructive manufacturing design test in which all of the units are stressed until failure, see Figure 1-1, similar to a post construction concrete quality sampling program where all of the core samples are quantitatively tested for strength using a destructive test method.



**Figure 1-1**  
**Complete Data**

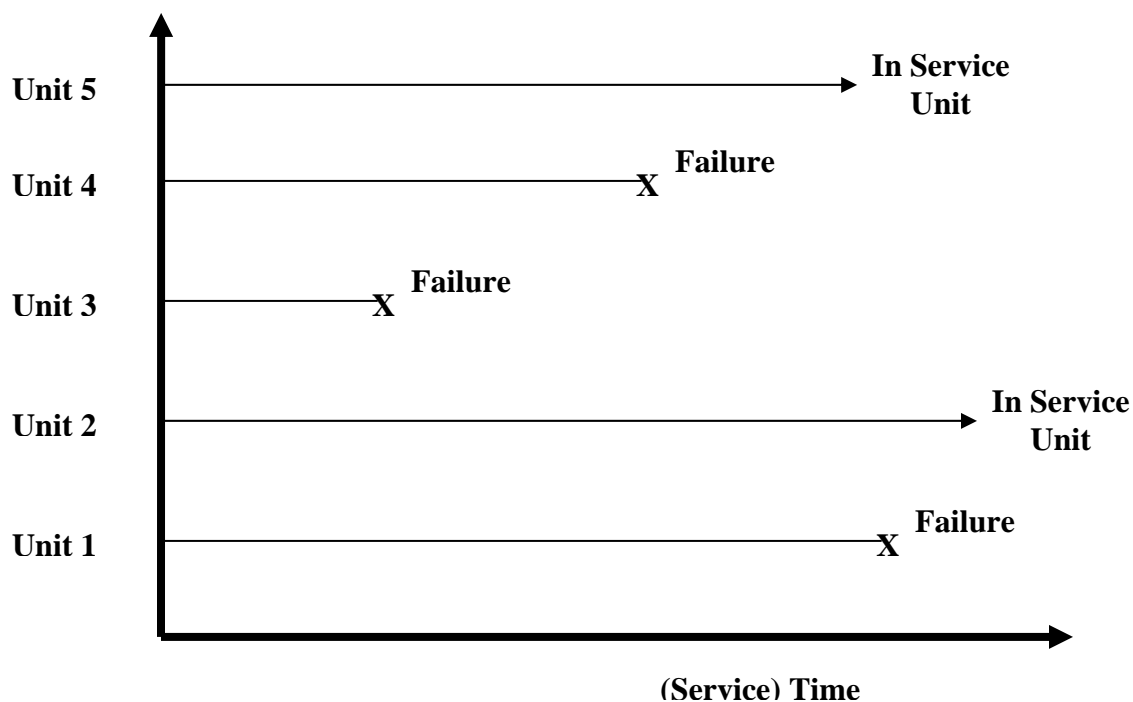
Most life data falls into the “incomplete data” classification. Examples include a manufacturing design test in which all of the units are operated under simulated life conditions for a maximum specified time, where some units survive and others fail within the test; the failure times are exact, but the test is terminated prior to the failing all of the units; the surviving units have an exact running time but lack an exact failure time, see Figure 1-2. Often referred to as “field data”, incomplete data is often associated with fleet or inspection data since many units have not failed (yet) or exact failure times are unknown, e.g., with inspection data the unit is known only to have failed sometime within an inspection interval.

## Censoring

With “incomplete data” statisticians describe surviving observations as “censored”. Censoring is a term that describes observations for which the exact failure time is unknown. Field data can have a very high percentage of censoring, e.g.,  $\geq 90\%$ , especially if the length of life is longer than the record retention or general population life, e.g., JMP’s sample Locomotive data is 61% censored and the life miles of the units are very high. In general most of the JMP sample data has a low percentage of censoring. As a result some of the practical questions that arise with highly censored data are not discussed in the help sections.

There are several general types of data censoring:

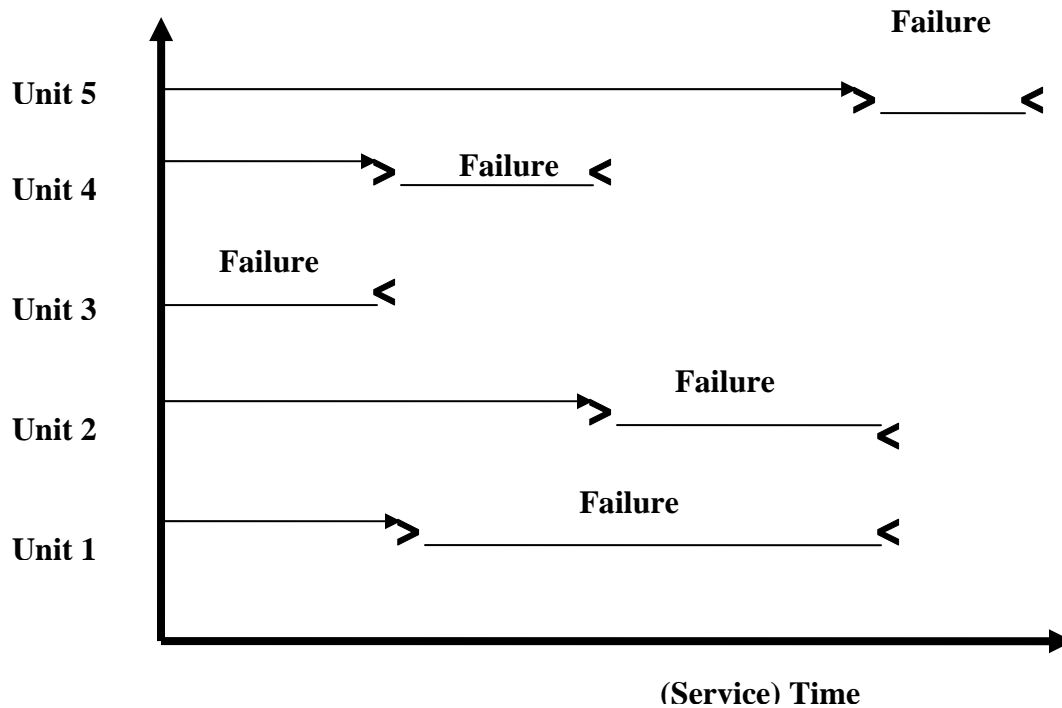
“Right censoring” refers to observations whose end state status is surviving at the end of the study or observation end, i.e., at the latest time recorded for the unit service segment. Since life data is generally displayed with an increasing time scale from left to right along the x axis, right censoring refers to the status on the right side of the graph. In Figure 1-2 the time axis is a relative service life axis, such that the all of the start times are aligned at zero, despite calendar time differences in real time.



**Figure 1-2**  
**Right Censored Data**

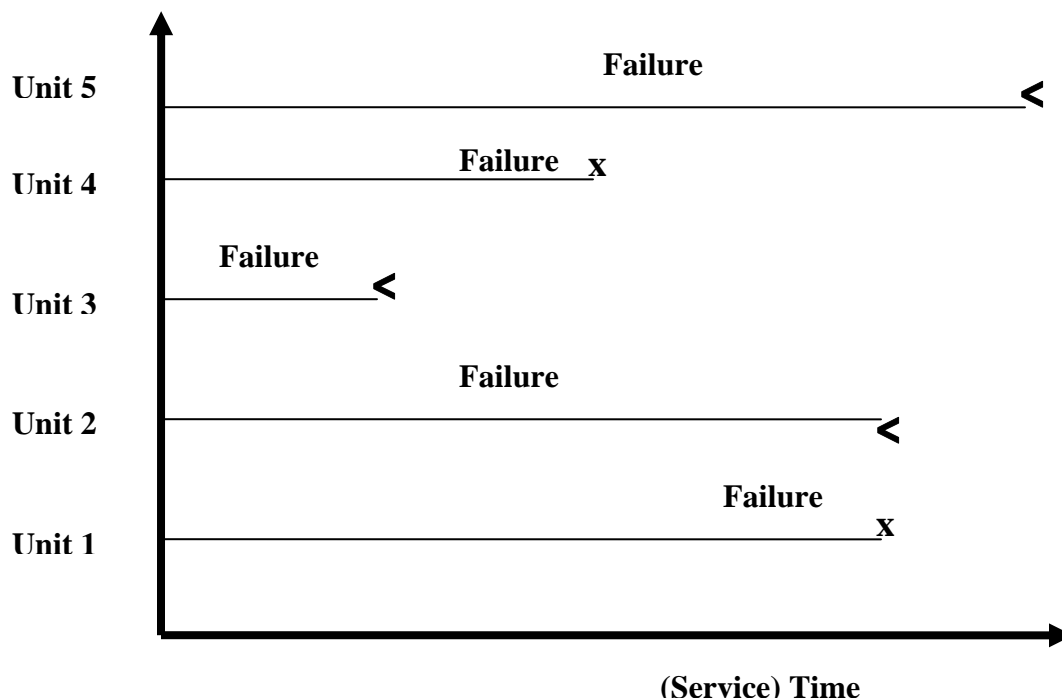
“Interval censoring” refers to uncertainty of exact failure times within an interval. Sometimes this is referred to as inspection censoring, since inspection data is an example of this type of data

where the failure time is not exact but known to occur between to known inspection dates, see Figure 1-3.



**Figure 1-3**  
**Interval Censored Data**

“Left censoring” refers to observations for which the failure is only known to occur to the left a known time/date. This type is a special form of interval censoring, where the starting time of the interval is zero. In other words one knows only that the failure occurred between zero and the left censoring time/date, see Figure 1-4.

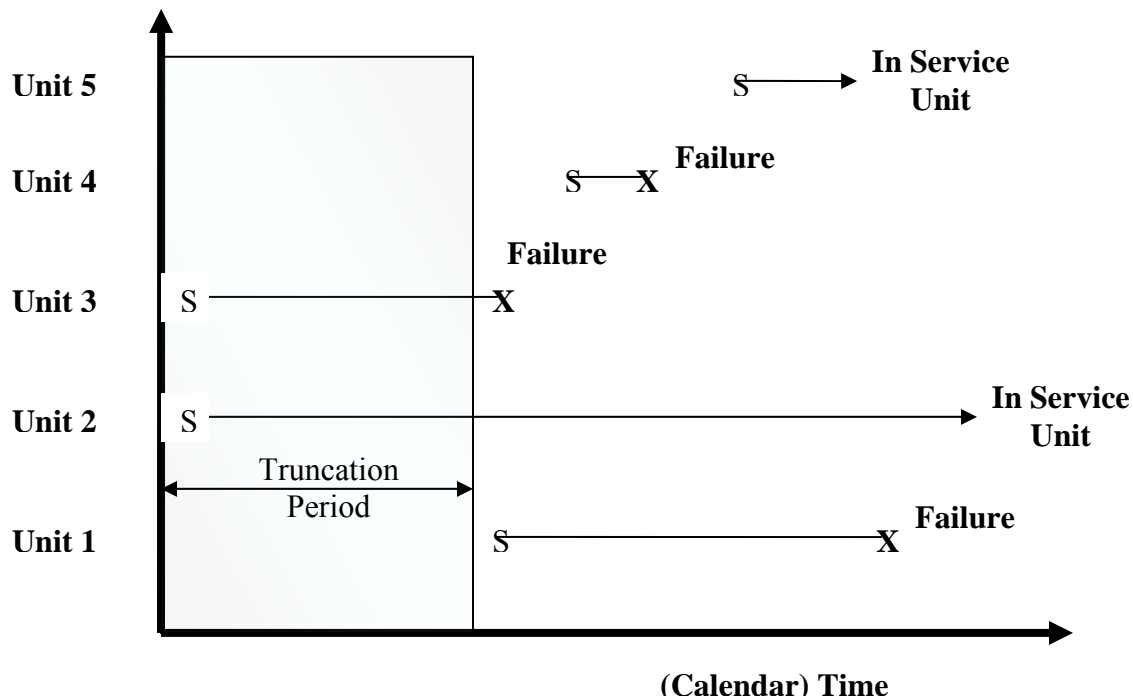


**Figure 1-4**  
**Left Censored Data**

Most life analysis methods have provisions for “incomplete data” and its censored observations. Some methods use the relative position of the observations without using the survival time data while other methods, e.g., use exact observation times, for failures and survivors, in the estimation of failure-time distributions and associated parameter estimates. As a result there are differences in the precision of the estimates between methods.

### **Truncation**

Truncation is a characteristic descriptor for data in which there are periods with known or suspected missing failure data. Truncation should not be confused with censoring. With censoring one knows the status (failure or survival) within the study, although the exact failure-time may be unknown, i.e., one knows the status for surviving units at the end date. With truncation, the status is unknown for a period within the study. For repairable systems with multiple segments the start/end time of the missing failure records are also be unknown, e.g., the data contains only the unit’s initial start time and last segment end time. In Figure 1-5, the time axis is shown on a calendar basis (as opposed to service time). Unit service segments are represented by the line segments starting with “S” and ending either with an “X” for failure, or an arrowhead for censored observations (surviving units). The early data period is highlighted by the box to represent truncation.



**Figure 1-5**  
**Truncated Data**

Suspected missing failures are discoverable when gaps in failure observations appear inconsistent with subject knowledge, due to lapses in failure archival data. Known missing failures are discoverable from conflicting data, e.g., when repairable system records indicate that a unit has had (at least) one previous failure (perhaps noted in comment fields or origination fields) but the data does not contain unit failure observations. One might view the latter case as left censoring, but with repairable systems multiple failures may have occurred in the truncation period and thus the left censoring date is indeterminable.

Most commercial statistical computer programs do not have the capability to handle truncation; however closed form equations can express truncation for computation by numerical methods, e.g., maximum likelihood estimation. The most powerful allies in the detection of truncation are good working knowledge of the subject's reliability and sufficient raw data review.

For example if service exposure is the process assumed constant and described adequately by a bathtub curve, then a gap in infant failures or a significant period with zero failures may provide the clues to data truncation identification. One can use JMP tools to visualize truncation using the two aforementioned allies. It would be helpful to have future JMP tools for user simulated truncation dates to assess parameter estimate precision and variability due to the extent of missing failure data or estimated truncation levels.

## **Check Assumption Validity with Subject Knowledge**

There are several underlying analysis assumptions regarding subject characteristics and data censoring which must be valid for the subject (physical specimen or process analyzed). Often one has knowledge of the subject's reliability and familiarity with the subject's usage application to know intuitively whether the analysis assumptions are valid. Other analysis assumptions may not be as intuitively obvious.

For instance, generally one knows (or can find out from subject matter experts) whether the subject component or system is repairable. On the other hand, it may not be obvious that failure modes are independent. In general the rule of thumb is to proceed cautiously with questionable applicability of underlying assumptions. The following section describes some of the key assumptions and subject knowledge areas that should be considered prior to analysis.

### ***Repairable vs. Non-Repairable Systems***

You need to know whether the subject is repairable or not. Does the unit have multiple records in the raw data? Could this be a sign of repairability? Or is it a sign of poor quality data? It is best to consult the subject matter experts if you don't know. One should consider this attribute carefully, since the appropriate analysis differs based upon this attribute.

### **Distinguishing 1<sup>st</sup> Failures of Repairable Systems**

For repairable systems 1<sup>st</sup> failures should be distinguished for separate analysis from repaired units since the repair processes can influence subsequent failures. If the mission of the analysis is to assess the total expected failures of the fleet, regardless of repair status for resource planning, then grouping repaired and new service histories may be appropriate. However, if the mission is to determine the most reliable units as supplied by the manufacturer, then the 1<sup>st</sup> service segments should be assessed separately, i.e., include units up to and including the 1<sup>st</sup> failure event. If the mission is to assess the efficacy of unit repairs then the 1<sup>st</sup> service segments and repaired unit segments should be separate subgroups, e.g., survival analysis of new vs. repaired units is a good starting point.

### ***Censoring Must Be Non-Informative***

For valid analyses for the analysis of censored life data, one must have censoring which is non-informative. This means that it is arbitrary with respect to the failure-time distribution. This could occur by fixed interval data collection, or by random selection. If however the units were removed from the study, due to a pre-cursor of failure (lead-warning indicator of failure) from a condition monitoring system, then the censoring scheme would violate this assumption and bias the analysis of the failure-time distribution.

## ***Independent Failure Modes***

There are subjects with multiple modes of failure throughout their lifetimes, e.g., infant mortality failures vs. random mid-life stress failures vs. late life wear-out failures. For JMP's Competing Cause analysis platform, these failures must be independent, i.e., one failure mode does not increase the probability of failure of another failure mode. Sometimes cause independence may be known or postulated by subject matter experts through working knowledge of service history and post mortem forensic analysis. Analysis platforms like the Competing Cause assume independence. One can assume validity when working knowledge corroborates independence to permit the use of Competing Cause analysis. The platform allows one to assess the projected lifetime extension through cause omission (postulated ability to remove specific failure modes).

## ***Underlying Failure-Time Distribution Known vs. Unknown***

Often one has a general knowledge of the subject's failure-time distribution without knowing the specific distribution, e.g., one knows there is a general bathtub shape but not the Weibull parameters for either end of product reliability. Other times one has a general sense that the failure time distribution is not age related (exponential distribution) without knowing a scale parameter. Still other times the subject is unknown and assumptions of the distribution are best left off in the initial analyses. In general since parametric tools have more power to quantify behavior, those tools are useful to test the appropriateness of the model. In some cases one may have a very good understanding of the shape parameter and insertion of it using Bayesian analysis, which is parametric analysis for the scale parameter with prior shape parameter knowledge.

## **Selection of Appropriate Statistical Tools**

There are several statistical tools at your disposal with JMP. Most life analysis tools fall within two main classifications: non parametric and parametric methods. Proper selection requires that one consider how well you know the subject that you are going to analyze. If you know nothing, you must approach with the least amount of assumptions, so non-parametric methods are a good starting point. If however the subject's reliability is well documented and studied one could focus more on parametric methods with due caution for data quality and specific batch, usage factors, underlying attribute, or process differences.

## **Chapter Summary**

The above discussion illustrates that reliability analysis is a powerful and somewhat specialized field of statistical analysis. JMP provides a multitude of tools to assist the analyst with discovery and visualization, and the importance of a strategic approach cannot be underestimated. Adopting a strategic method to reliability analysis is a good way to ensure a successful mission and working efficiency





## 2 CHAPTER 2

# TACTICAL OBJECTIVES: JMP<sup>®</sup> SEQUENCE

---

A great thing about JMP is its ability to enable roaming discovery. With JMP an analyst can select any order of analysis as patterns emerge and unravel progressively through sequential analysis. Most of us prefer this like the wanderlust of unfettered travel. However, there are some recommended paths of JMP analysis that follow a strategic approach to life data principles and assumptions and enable efficient discovery. As discussed in previous sections, it is best to start with the least amount of preconceptions about the data and the subject's reliability. Non-parametric methods are suited to this approach and include analysis platforms that are both within and outside of JMP's **Reliability and Survival** platform. For users unfamiliar with life data concepts and JMP platforms, initial data exploration with traditional JMP platforms, e.g., **Distribution**, **ANOVA**, etc., can provide a good introduction to the data and reinforce later inferences through discovery with JMP's **Reliability and Survival** tools.

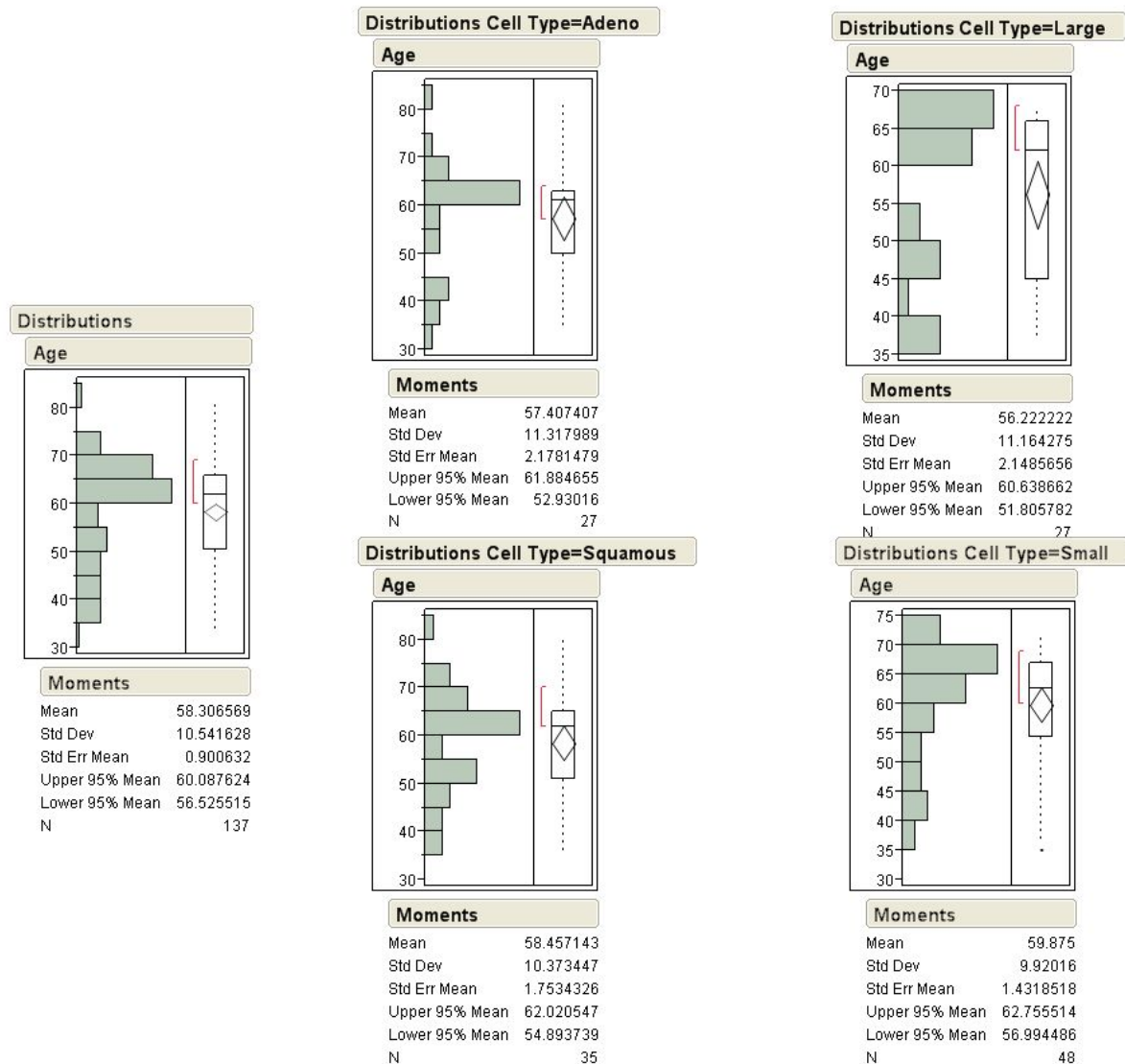
### Distribution Analysis Platform

This is always a good place to start the non-parametric analysis. JMP 8 adds the **Graph Builder** platform which could be used as an alternative given its intuitive feel, but the data table selection and highlighting features with this platform are still an advantage of this platform. The histogram display is intuitive to statistically oriented folks but non-stat folks are more familiar with pie charts. Tests and fits for normality are built into this platform but have limited applicability with life data. This platform is also a good summary statistics tool for initial data review and cleanup, e.g., **Recode** features within the **Data Table** platform.

To begin with this platform, select **Distribution** from the JMP menu and analyze several distributions of the time domain fields, e.g., end state status field, failure modes, and categorical descriptor fields.

Useful inferences from this might include whether the fleet has age centrality, e.g., near normal distribution, or perhaps fleet age is uniformly distributed. Depending on the failure-time distribution this may infer that the fleet is at a decreasing infant mortality risk or increasing wear-out risk.

In Figure 2-1, the VA Lung Cancer JMP sample data (distribution of Age and distribution of Age by cell type) there appears to be some centrality of lung cancer at age 60-70. When subdivided by Censor (Survival vs. Mortality) the mortality centrality is further reinforced at the Age 60-70, whereas the survivors are at a lower age. When subdivided by Cell Type, the Adeno and Squamos look to have more age centrality than the other cell types. This perhaps suggests a different life process.

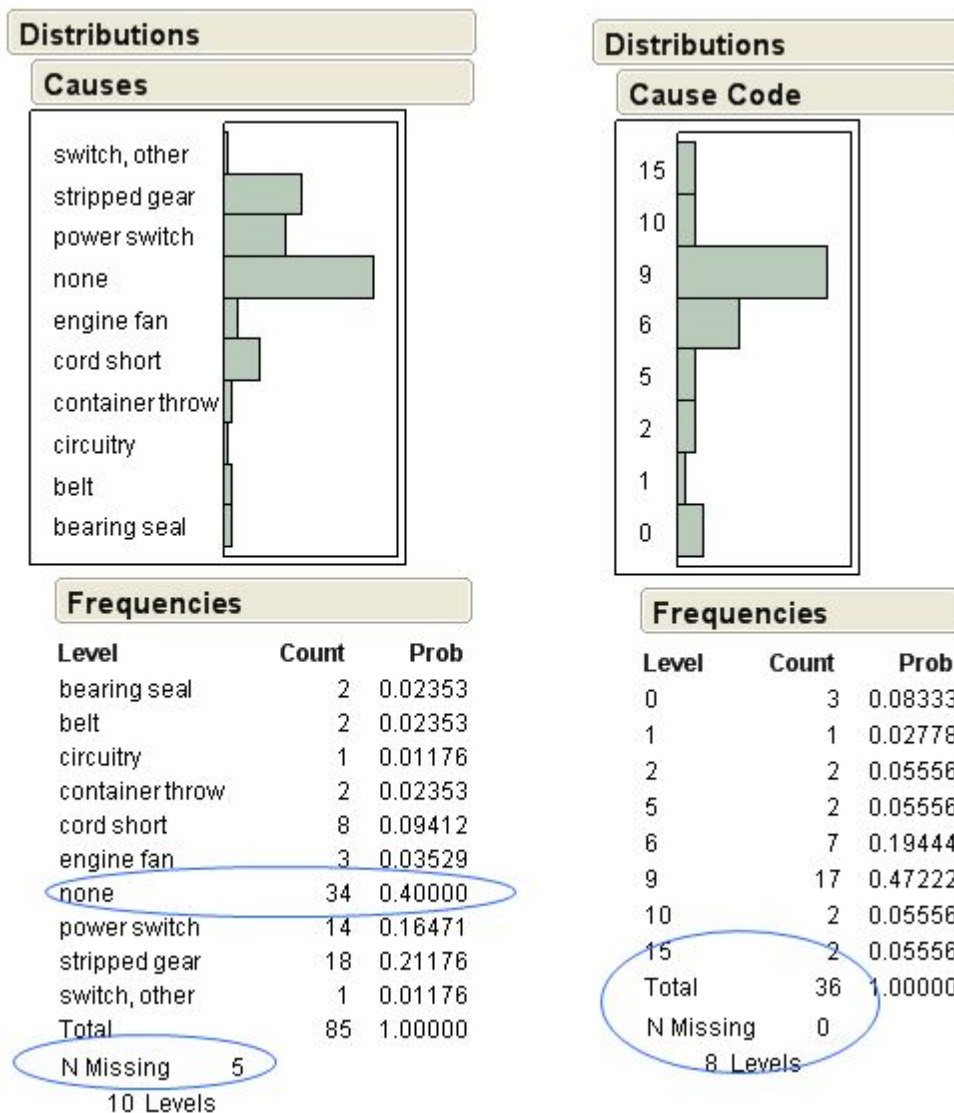


**Figure 2-1**  
**Distribution Analysis**

Data gaps are easily detected when the distributions have significant Null populations in the distributions, e.g., distributions of start date, end date, or elapsed time to failure fields. This is most easily seen by reviewing the “**N Missing**” in the **Distribution** platform. In the Blender example, Figure 2-2, (left distribution) “none” is the most predominant cause, which is sometimes the case with field data. Data formatting issues are also obvious from distributions of categorical descriptions and cause description fields. The selection and highlighting features can be used, e.g., with the **Recode** command.

In Figure 2-2, Blenders, in JMP sample data (Cause = “none” or NULL; Compare to Appliance JMP sample data for Cause Code where there are no NULL or formatting discrepancies)

illustrates the use of **Distribution** platform for initial data reviews in life data analysis. {The reader can try entering in a year typo like 2999 instead of 1999 or a misspelling/formatting to use **Recode** feature}.



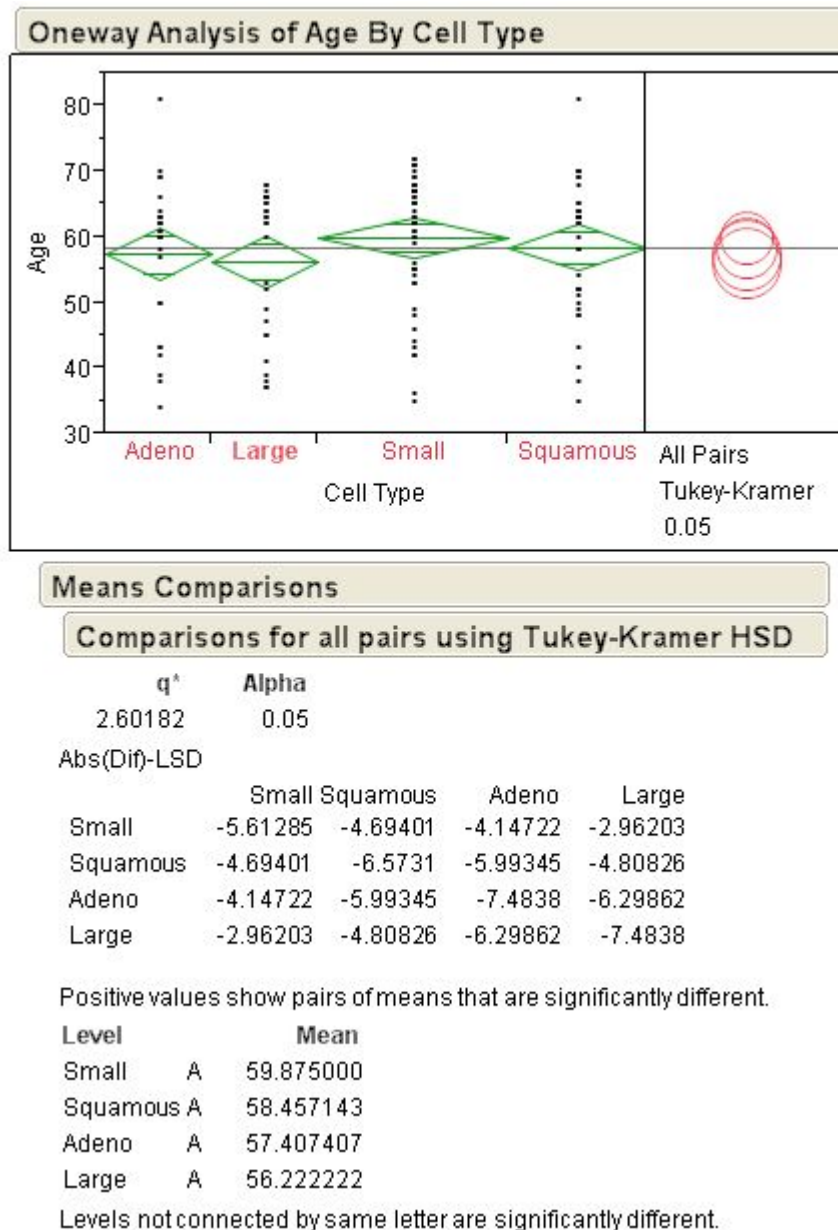
**Figure 2-2**  
**Distribution Analysis for Initial Data Quality**

## ANOVA Analysis Platform

Analysis of variance (ANOVA) is another non-parametric tool that helps visualize differences by categorical factors. For statisticians it is intuitive and the plot markup features and difference tests provide quick interpretation of the results. Non-statistical folks may need some explanation to interpret the meaning of the groups of distributed data points. With a little explanation however subject matter experts can generally assess the results as consistent with subject knowledge (or smell a rat).

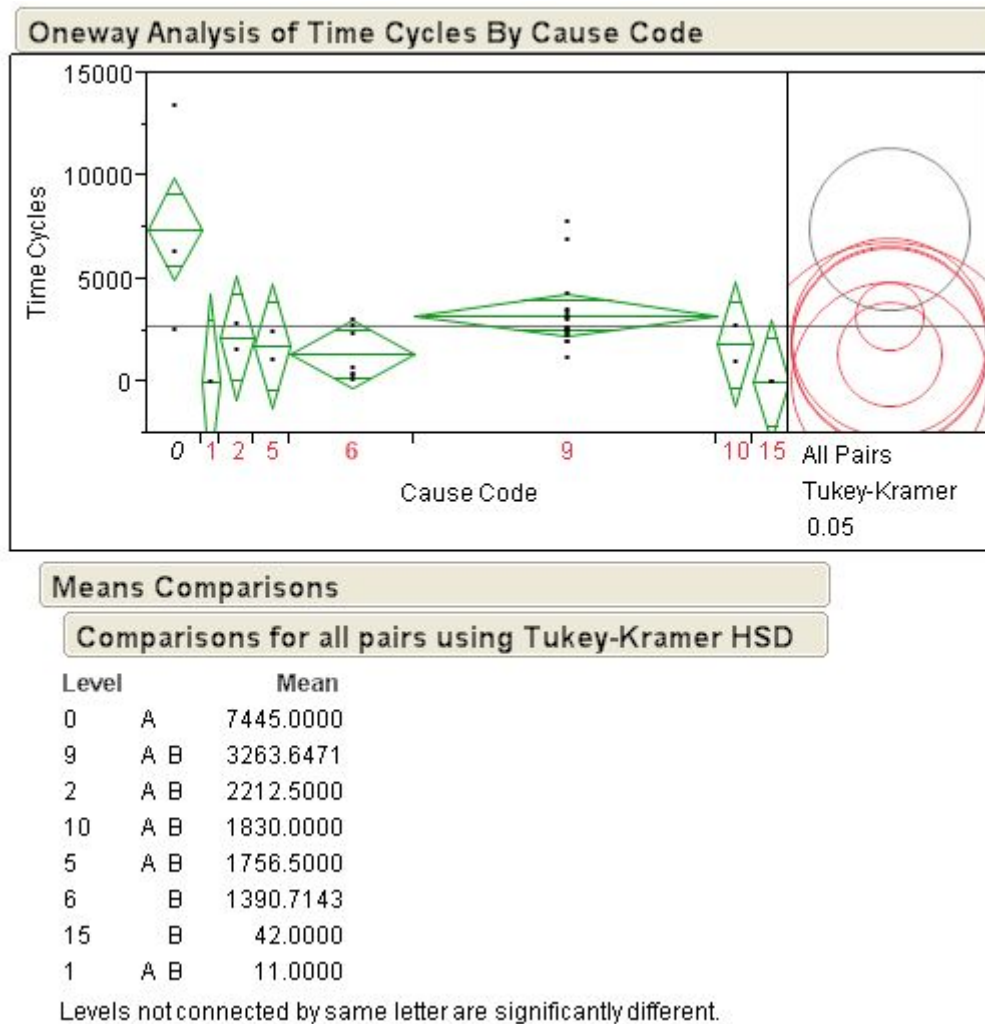
To use this platform select **Fit Y by X** platform from the JMP menu and analyze the categorical group differences, e.g., by start and end dates/times (y) by end state status (x), by age (y) by end state status (x), by age (y) by failure mode/cause (x).

In Figure 2-3, Lung Cancer JMP sample data illustrates the further exploration of differences cited in the distribution – different but not significant – and the corresponding Means tests.



**Figure 2-3**  
**ANOVA Example without Significant Mean Difference**

In Figure 2-4, the Appliance JMP sample data illustrates a means test of significant difference by cause code.



**Figure 2-4**  
**ANOVA Example with Significant Mean Difference**

## Graph Builder Analysis Platform

Visually explore the data for the same distribution and difference inferences by using the **Graph Builder** features of JMP 8. Select **Graph**, then **Graph Builder** from the JMP menu and begin by selecting time fields for the x axis. Explore the effects of adding additional variables. It's like a visual "Pivot table" format (in Microsoft's Excel®). Don't be afraid to try several combinations, some of the best discovery insights are "accidental". The fit line (right click on plot and select **Add Line**) is more of a trend line, and can bring interpretation questions upon presentation.

In Figure 2-5, Blenders JMP sample data (time cycles on x; cause on group by y; overlay manual/auto) illustrates the use of **Graph Builder** platform to analyze life data.

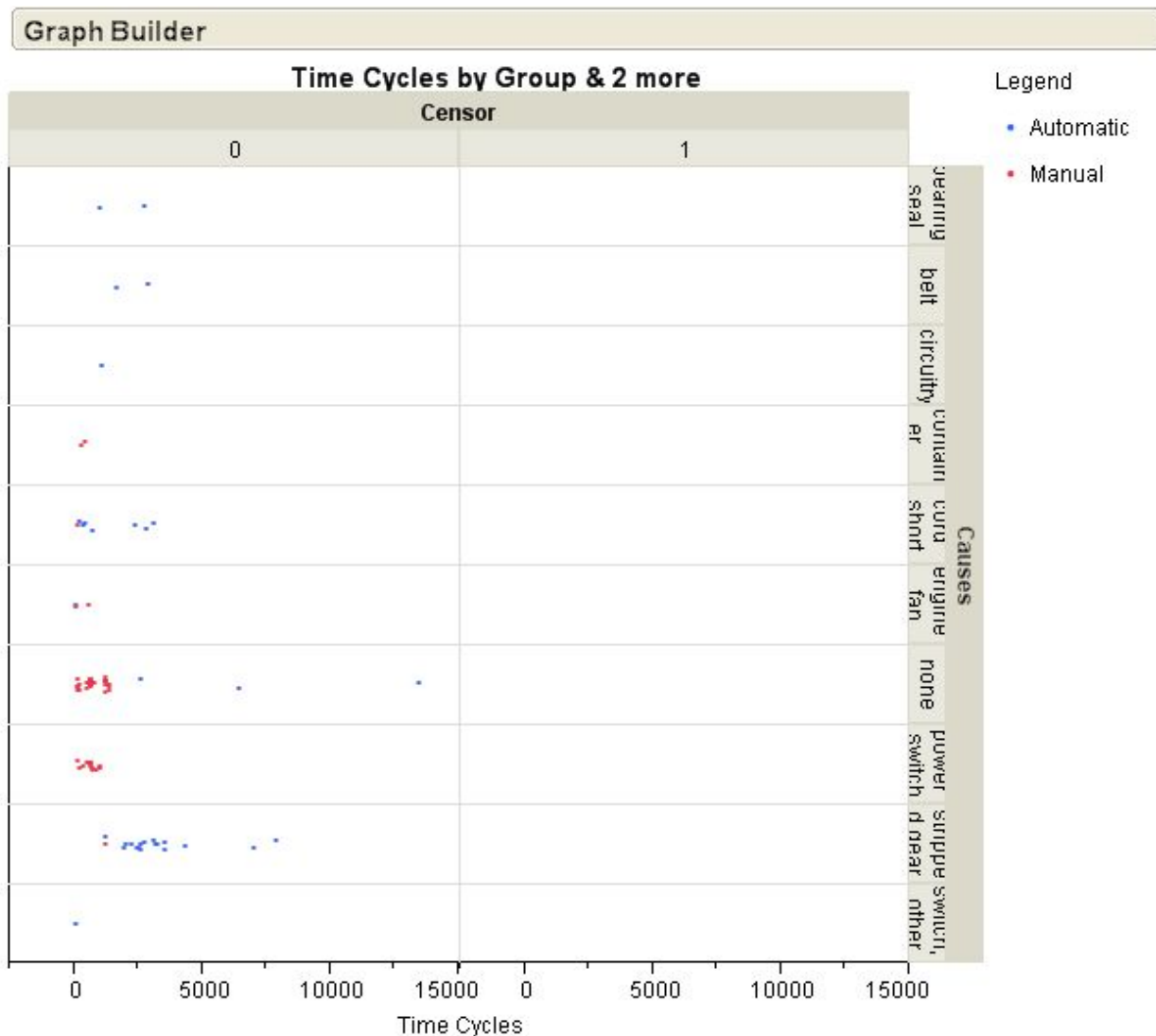


Figure 2-5  
Graph Builder

## Tabulate Analysis Platform

Explore the **Tabulate** options to summarize data in tabular format. This may provide a more familiar output for less visually oriented audiences. Although the results are the same as the **Graph Builder**, the output may require less interpretation by others and less textual explanation in formal reports. **Tabulate** also has features to summarize the NULL cells which can be used to inspect data quality or communicate underlying data gaps that prevent some analyses.

Use **Tables>Tabulate** from JMP menu:

Figure 2-6, VA Lung Cancer JMP Sample data (Cell Type Column >Treatment underneath <> KPS Grouping column (row) <> try various fields to column and row) illustrates the use of **Tabulate** to analyze life data.

Tabulate								
KPS	Cell Type							
	Adeno		Large		Small		Squamous	
	Treatment		Treatment		Treatment		Treatment	
	Standard	Test	Standard	Test	Standard	Test	Standard	Test
10	0	1	0	0	0	0	0	0
20	1	0	0	0	1	2	1	2
30	1	0	0	3	4	4	1	1
40	1	5	1	0	5	2	2	0
50	1	2	2	0	3	1	1	4
60	0	4	2	4	9	2	4	2
70	1	2	3	2	2	5	4	4
75	0	0	0	1	1	0	0	0
80	4	2	5	2	5	1	2	3
85	0	0	0	0	0	1	0	0
90	0	1	2	0	0	0	0	4
99	0	1	0	0	0	0	0	0
All	9	18	15	12	30	18	15	20

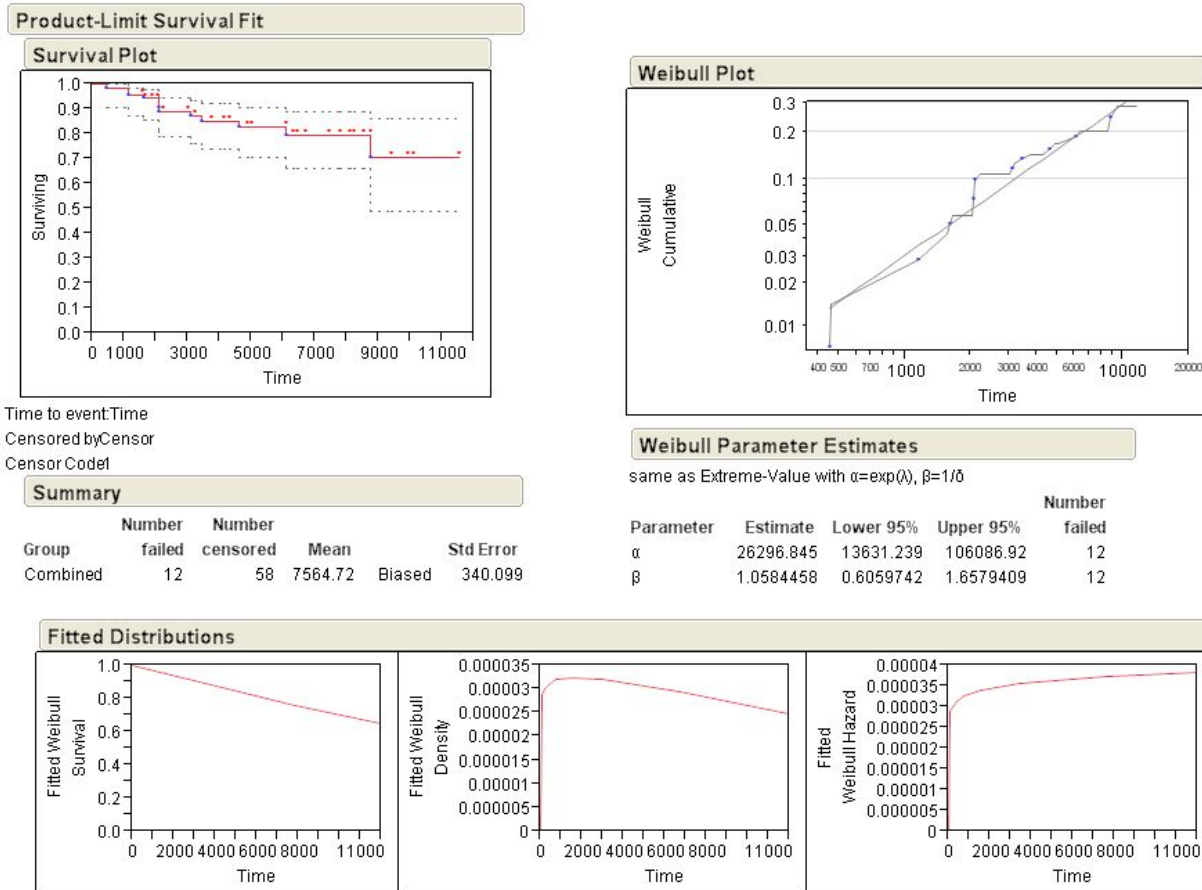
Figure 2-6  
Tabulate

## Survival Analysis Platform

Use the **Survival** Platform to assess the non parametric distribution. One can also use the **Life Distribution Platform** to accomplish the same thing by choosing the non parametric fit option. The cumulative fitted **Hazard Plot** is available as an option. Select **Analyze>Reliability & Survival >Survival** from the JMP menu:

In Figure 2-7, Fan from the JMP Sample data (Time as time to failure, Censor = Censor, Add Points in Plot Options and Confidence bounds. Go back to table and select Rows> Color or Mark by Column> Select Censor Column) illustrates the surviving and failing units on the plot and the mid step quantiles. Select **Weibull** plot and fit and >**Fitted Plots** for a fitted Weibull Hazard plot.





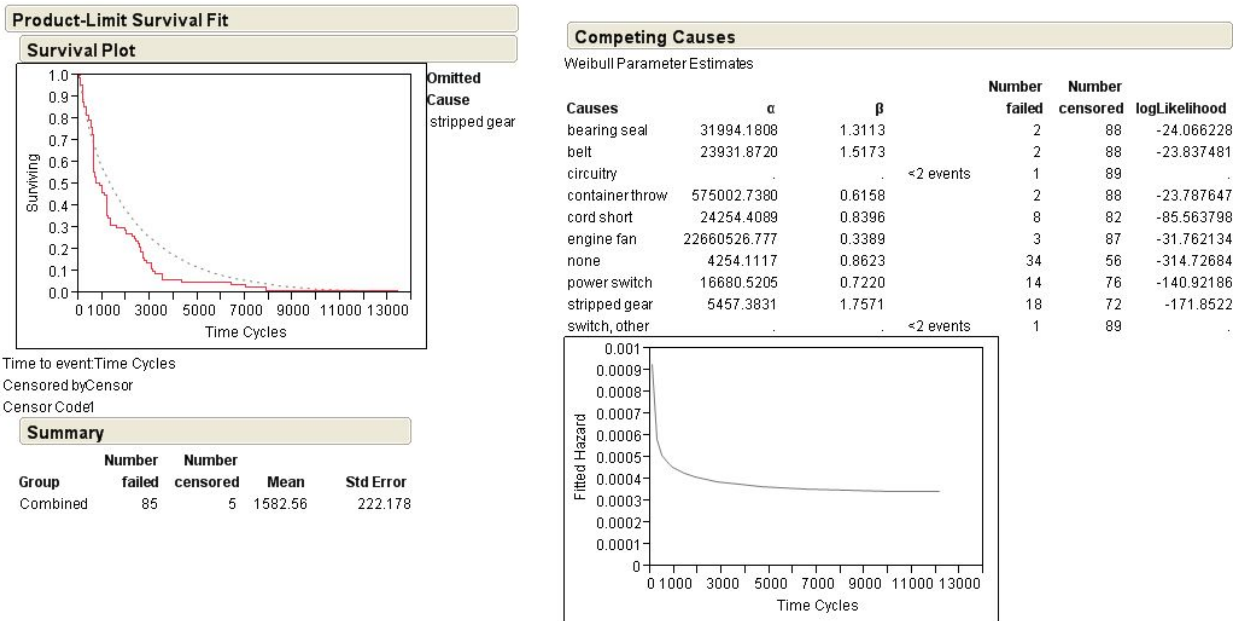
**Figure 2-7**  
**Survival Analysis**

## Competing Cause Analysis

This option permits one to assess the failure causes (modes) when the assumption of cause independence is valid, to see the hazard contribution of the predominant failure modes. Some simulation capabilities are provided within this platform to generate simulated failure data set (complete data) for the calculated Weibull parameters calculated from the original data fit.

In Figure 2-8, Blenders JMP Sample Data (time cycles = time to failure; Censor = Censor; Select Competing Cause from Survival Plot options, then select Cause = Causes; then Select from Competing Causes >Hazard Plot> then also >Omit Causes, power switch or stripped gear; Watch resulting change in Hazard Plot and dashed Survival plot with omission) illustrates use of Competing Cause analysis on life data.



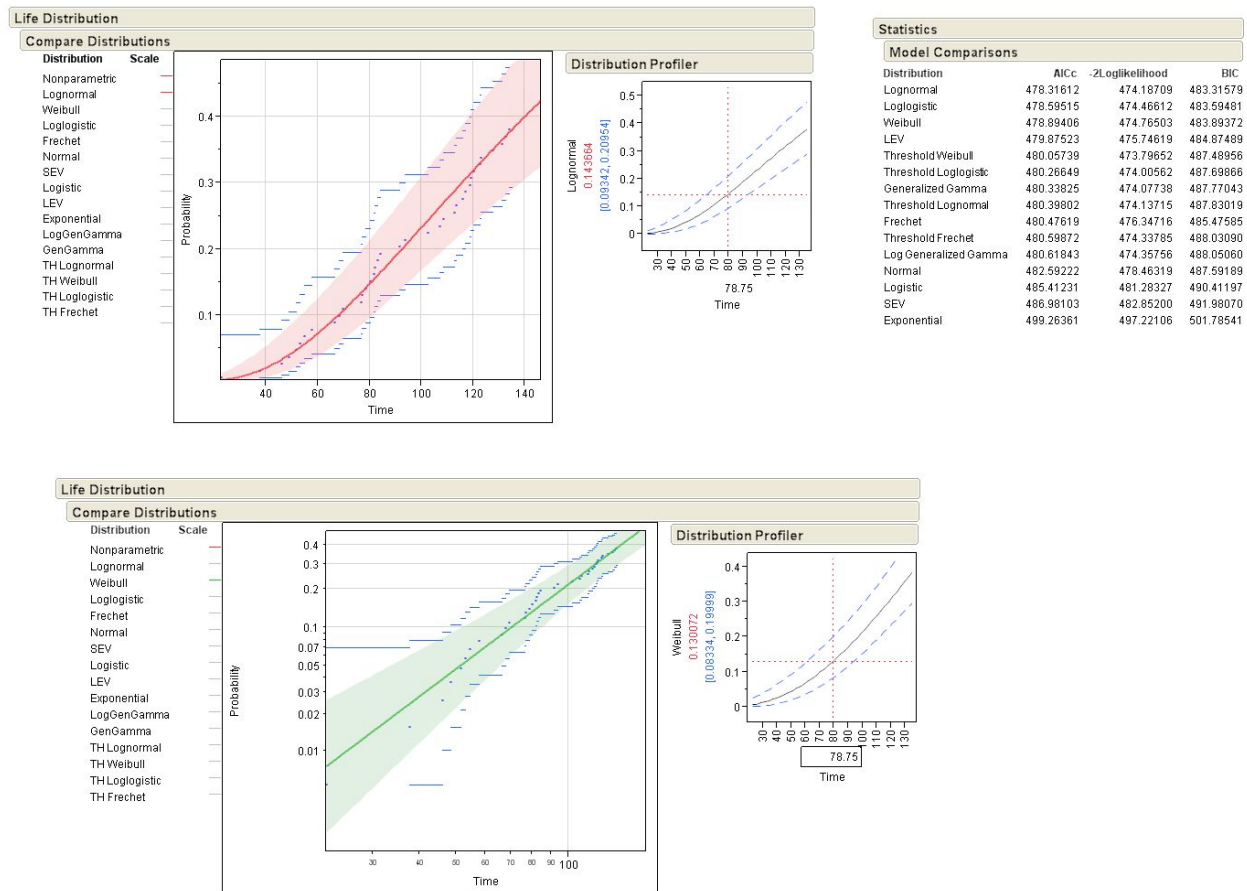


**Figure 2-8**  
**Competing Cause Analysis**

## Life Distribution Platform

This new platform to JMP 8 allows users to compare the non parametric fit to many parametric distributions. The platform is very graphical and intuitive with user driven distribution selection. Results are ranked by log likelihood values for “best fit”. Use caution since the highest ranking models are typically “over-fit”, i.e., more parameters than one can readily attain, explain, or expect to reproduce in the future process behavior (so KISS [keep it simple stupid] is especially applicable here). See how big the model difference is from a simple two parameter model before using a more complex distribution.

In Figure 2-9, Locomotive JMP Sample Data (Failure plot default display; select LogNormal, then Change scale to LogNormal display - changes axis for fit line; see parameters, then select Weibull and Weibull scale for overlay and axis change. Compare parameters. Now from Life Distribution menu select> Fit All Distributions then review all Model comparisons) illustrates the use of JMP’s Life Distribution analysis on life data.



**Figure 2-9**  
**Life Distribution**

## Fit Life by X

This new JMP 8 platform permits one to analyze life data when only one (failure) factor is present. It is used for accelerated aging tests and usage/exposure factors in life data modeling. If the factor data is available from field data or by design of experiment, it is a powerful tool to explore the effects of usage/exposure factors which accelerate failure. Good examples are shown in JMP help sections using the DeValt JMP Sample data, shown in Figure 2-10. The new platform allows one to compare the accelerating factor groups with several underlying life distributions.

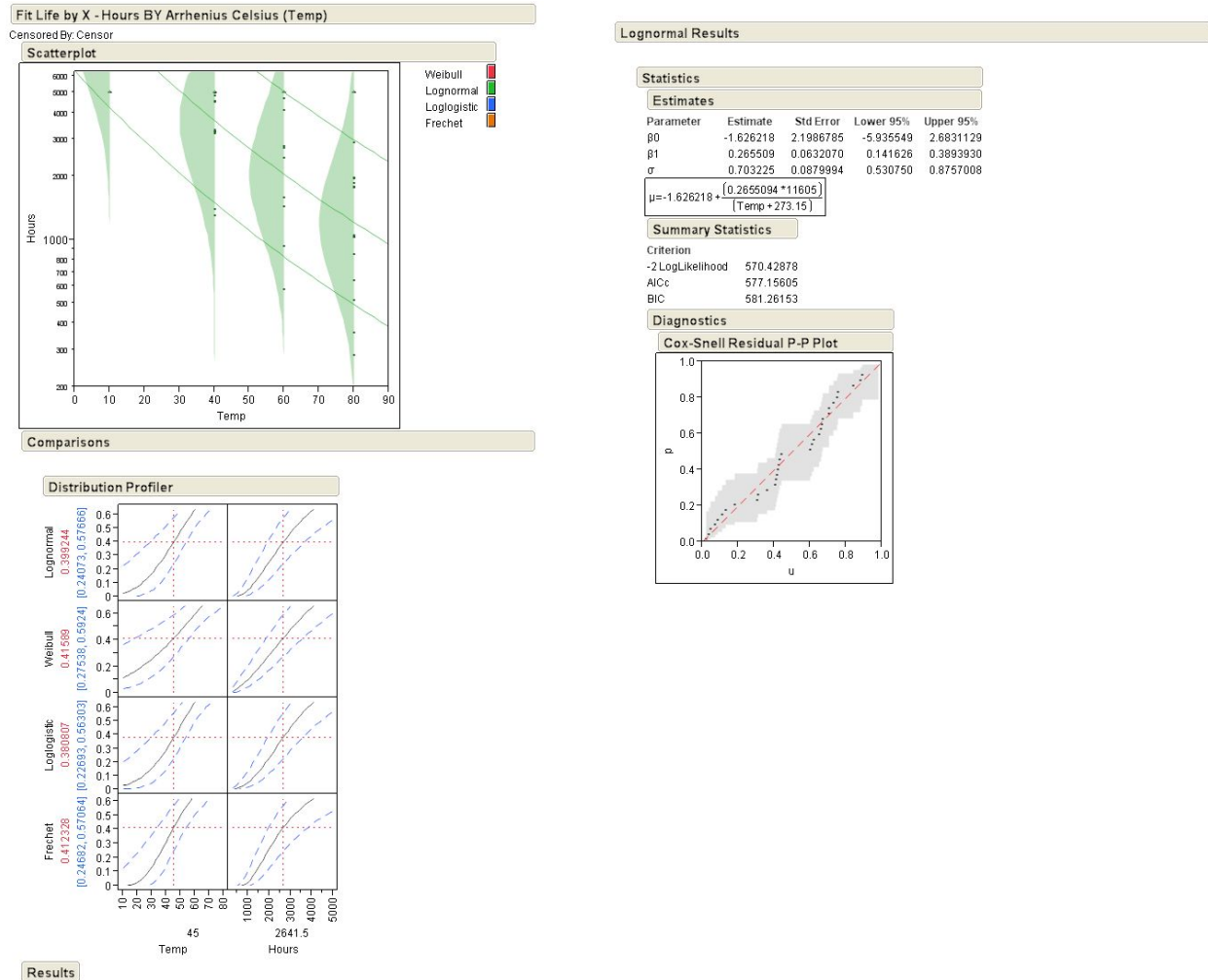


Figure 2-10  
Fit Life by X

## Recurrence Analysis

This platform is used for repairable systems or for life data where there are multiple events per unit. Typically the unit or patient has a sequential history of repairs or treatments for recurring conditions, e.g. the treatment of a long term disease. With field data the data quality may be spotty and leave gaps in the repair and return-to-service history, i.e., truncation. Recurrence analysis permits the analysis of repairable systems and the quantification of the mean cumulative function, which is cost per unit as a function of time. The cost can be the number of repairs or the actual repair cost. Good examples are shown in JMP help using the Valve Seat Repair JMP Sample Data.

## **Survival and Reliability Analysis II**

There are many more advanced analysis features within JMP but the detailed discussion of these capabilities is outside the scope of this paper/presentation. These features are useful for complex and specialized analysis needs.

## **Chapter Summary**

JMP 8's life analysis platform adds to the tools available to explore life data. This chapter provides a generalized sequence of JMP analyses as a tactical guide for efficient discovery, but variations are often necessary and may provide more efficient discovery and insight due to unique factors and skills of the analyst. Missions with more complexity and special data challenges, e.g., multiple modeling factors, limited failure observations, and semi-parametric modeling (discussed in JMP literature) are beyond the scope of this paper.

# 3

## CHAPTER 3

### PRESENTATION OF RESULTS

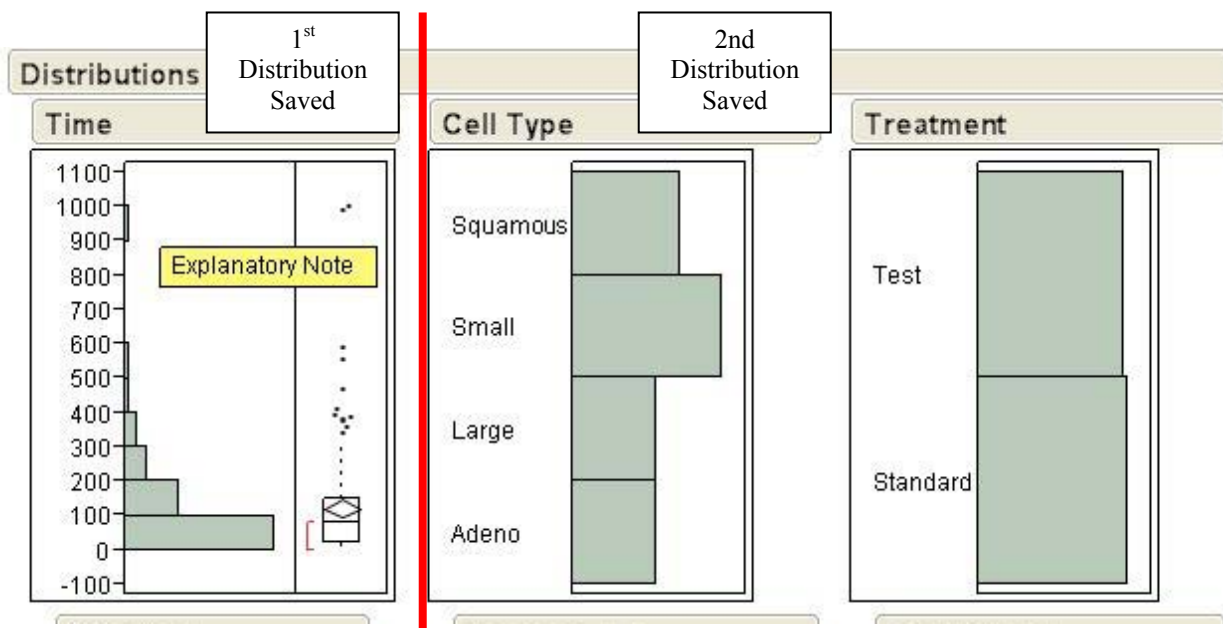
---

#### Using JMP® Features to Save Results

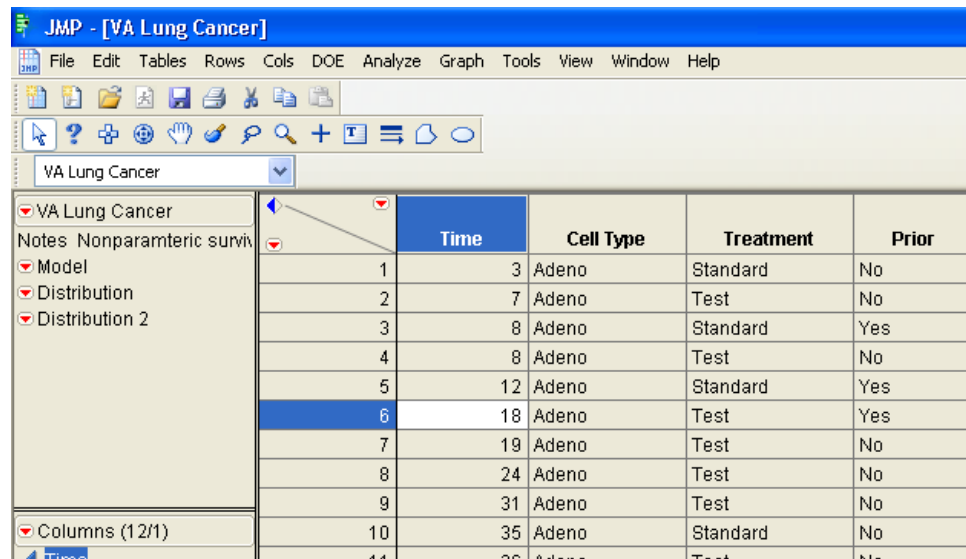
There are several options within JMP to save analysis results as you explore the data and conduct analysis. Use these to record your notes as well and striking results for later cut/paste to final report formats.

#### *Save to Table Scripts*

These are then named saved for later retrieval and electronic demonstration, which are good for interactive analysis. Note that explanatory notes (yellow label) as shown in Figure 3-1 are not saved in this format. It is a good practice to name the scripts with names that can be easily recalled and directed to the live demonstration, since the default convention is generic sequential naming, e.g., Distribution 1 as shown in Figure 3-2.



**Figure 3-1**  
**Save Results to Table**



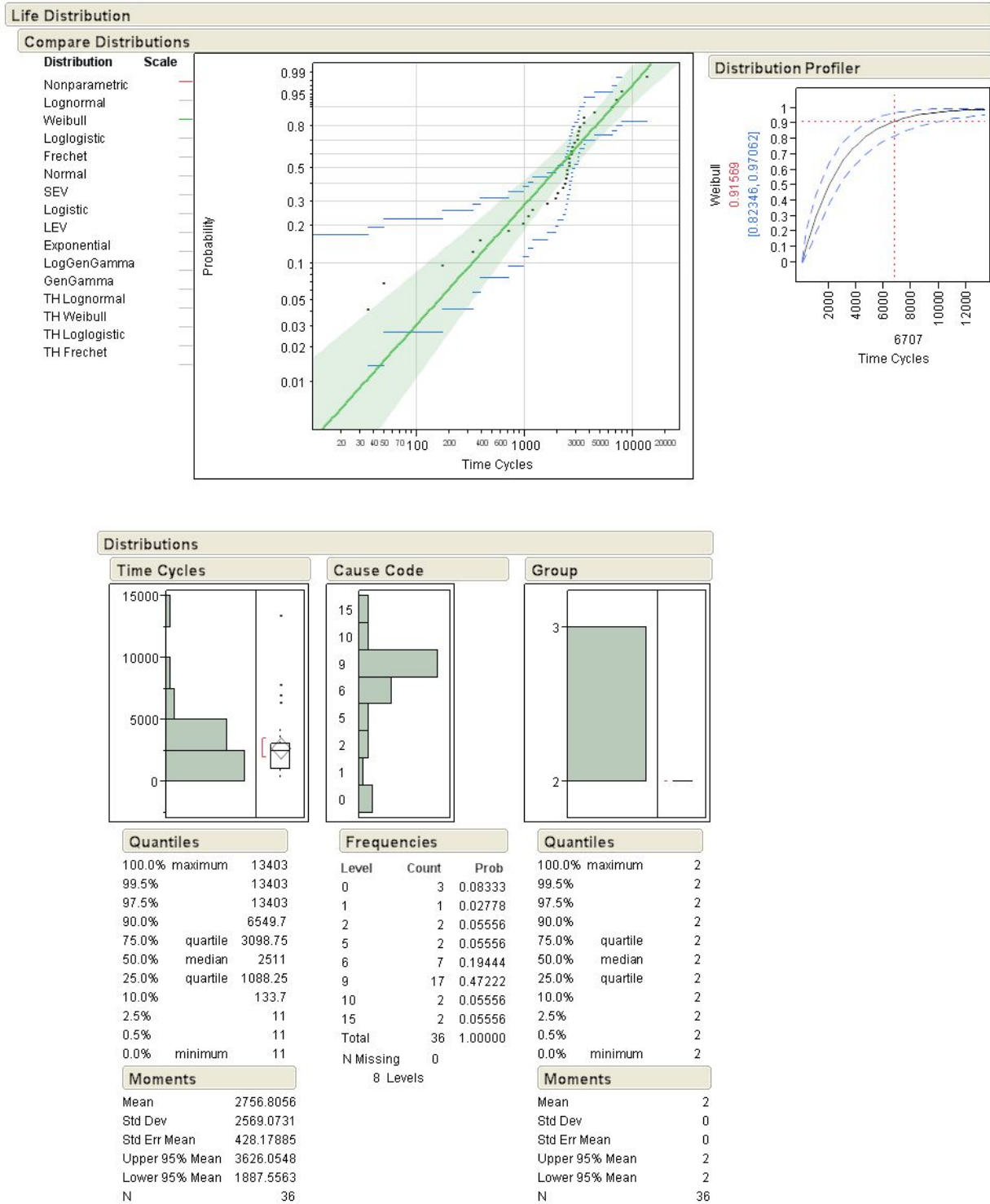
	Time	Cell Type	Treatment	Prior
1	3	Adeno	Standard	No
2	7	Adeno	Test	No
3	8	Adeno	Standard	Yes
4	8	Adeno	Test	No
5	12	Adeno	Standard	Yes
6	18	Adeno	Test	Yes
7	19	Adeno	Test	No
8	24	Adeno	Test	No
9	31	Adeno	Test	No
10	35	Adeno	Standard	No
11	36	Adeno	Test	No

**Figure 3-2**  
**Naming Saved Results in Table**

## Journals

These are good for pasting striking analyses snapshots as you go (incrementally appending snapshots during a session) for later cut and paste, but without the interactive capability later on. An advantage is that explanatory notes (graph labels) are saved in the snapshot which is useful for prepping JMP shots for slides, documents, etc. The **Layout** command allows one to manipulate the placement of images within the analysis for publishing and presentation layouts. These can be saved in many different output formats to accommodate cut and paste actions easily.

The sample **Journal** shown in Figure 3-3 illustrates how JMP results can be sequentially appended to the **Journal** in any user directed order by the analyst.



**Figure 3-3**  
**Save Results to Journal Example**

## Tune the Output to the Audience Preference When Possible

In general the amount of detail in JMP screen shots is overwhelming to novice viewers. Don't overload them too many new visualizations, or you'll spend too much time explaining the JMP displays and not the important analysis results, inferences, and application interpretations. Give them what they prefer if possible - statistical orientation, e.g., distributions, statistical tests, etc. vs. non-statistical, e.g., pie charts, graphs, tables, etc. The old adage KISS applies.

### Visual Output

Analytical audiences love graphs and charts and have experience interpreting them. Generally they prefer charts and graphs. This generally includes statisticians, engineers, and many managers too. Be prepared to explain all the details, aka, don't show them something you can't explain. For example most of the ANOVA analysis output is fairly intuitive, but the visual tests for difference may not be familiar to non-statisticians. Engineers will question "bad fits", so JMP results that use underlying MLE methods may receive scrutiny on goodness of fit.

In Figure 3-4, the MLE fit of a Weibull plot (Fans JMP sample data) that has data points off the line. The dashed line in Figure 3-4 is shown to illustrate where an audience might "eyeball" the fit line which simply fits only the failure observations. The MLE fit however accounts for the times of all of the observations (surviving and failed), which is not intuitively obvious.

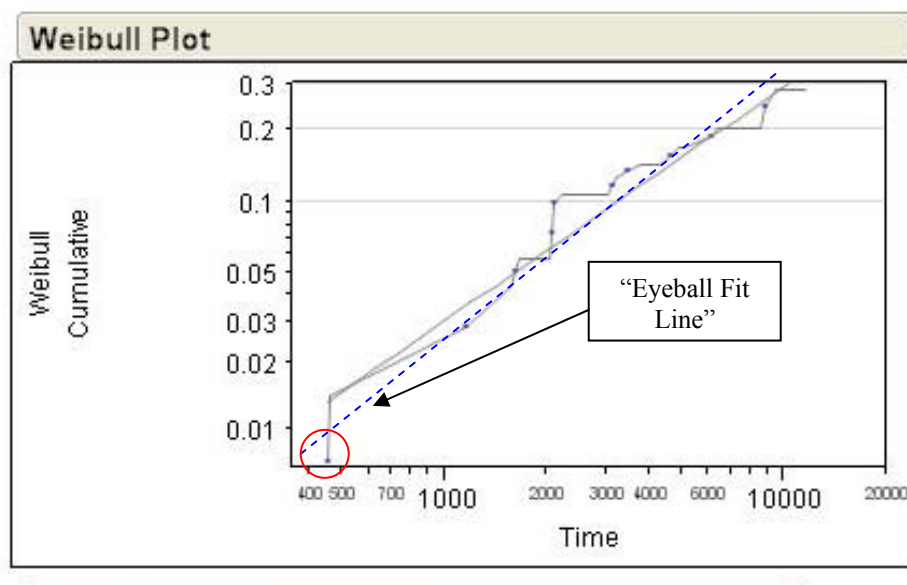
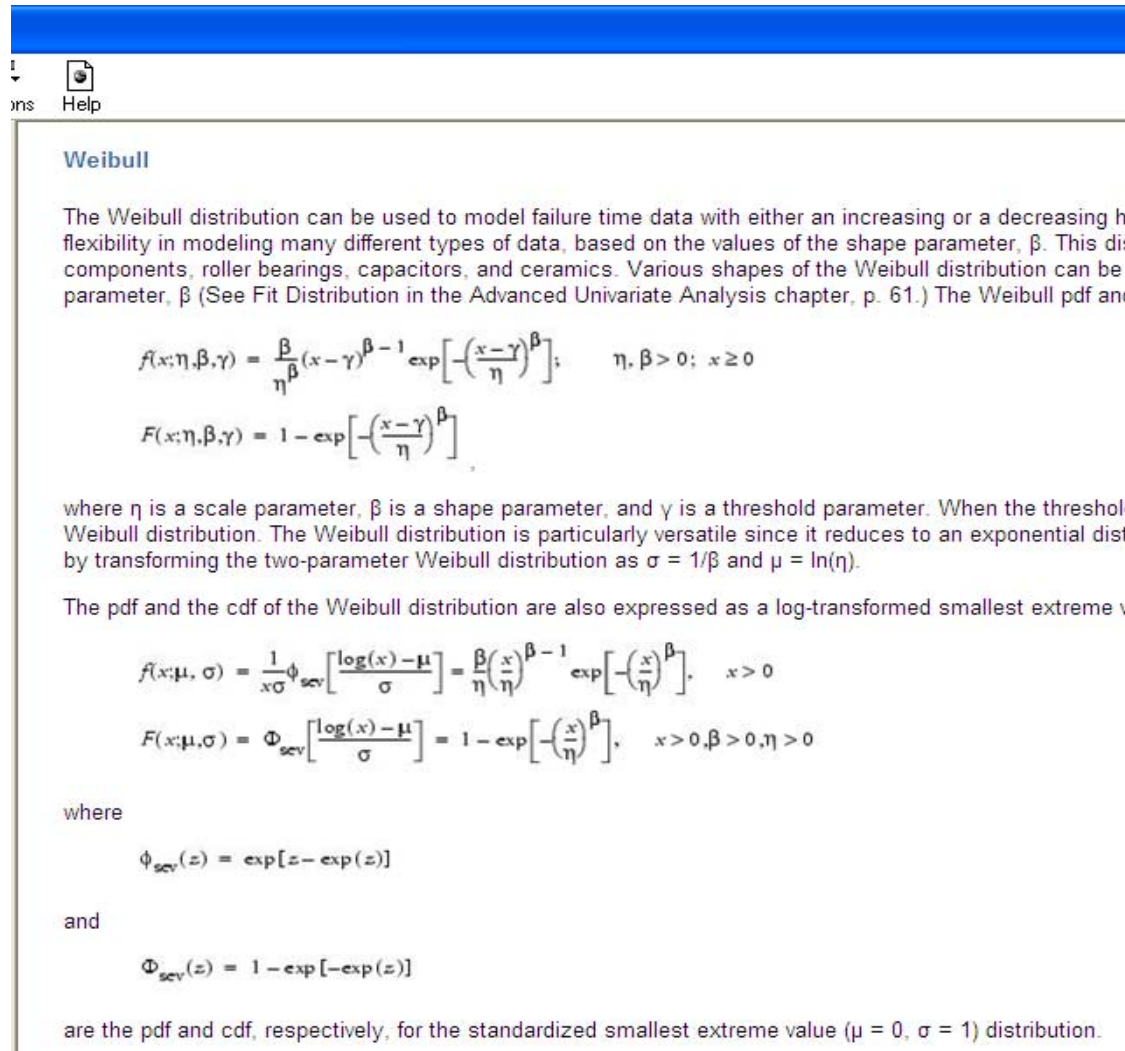


Figure 3-4  
Visual Output: MLE Fit Example



## Statistical Theory within Text Results

The world of statistics is bursting with its own symbols, vocabulary, and acronyms, as shown in Figure 3-5. For non-statisticians the terminology is a turn off. Many decision makers have a limited exposure (and tolerance) to the Greek alphabet and statistical conventions in text format. If such material must be presented, then provide definitions for all symbols. Choose your level of detail carefully and present only what you are prepared to explain!



**Figure 3-5**  
**Statistical Theory Presentation Example**



# 4 CHAPTER

## SUMMARY & CONCLUSIONS

---

### Summary

The discussion has illustrated the powerful tools available with JMP 8, but like all power tools, the real power depends on the skill of the user.

### Conclusions

The successful application of JMP in reliability analysis requires a strategic plan, proper tactical sequencing, and thoughtful presentation of results.

### Recommendations for Future JMP Features

Some of the additional features that would improve JMP's functionality include:

- Truncation functionality (create the ability to interactively set truncation periods across data sets and view the resultant parameter estimates),
- Monte Carlo simulation capability expansion (Like Weibull ++® 's QCP feature), including simulated data set generation for user specified parameter values and adjustable levels of censoring to assess the comparison of variability by MRR vs. MLE methods,
- MRR method plotting options (regression on y or x) capability and manual plotting paper slope tools,
- Point labeling capability on survival plots that show points,
- Retention of explanatory labels on analysis results when saved by script method to data table.



# A

## APPENDIX: BIBLIOGRAPHY

---

1. William Q. Meeker, Luis A. Escobar, *Statistical Methods for Reliability Data*, Wiley, New York, 1998
2. Wayne B. Nelson, *Applied Life Data Analysis*, Wiley, New Jersey, 1982
3. Robert B. Abernathy, *The New Weibull Handbook*, Florida, 2004
4. JMP 8.0.1® User Guides, electronic program documentation
5. WEIBULL ++7® Program *Life Data Analysis Reference*, ReliaSoft Corporation Publishing, 2008