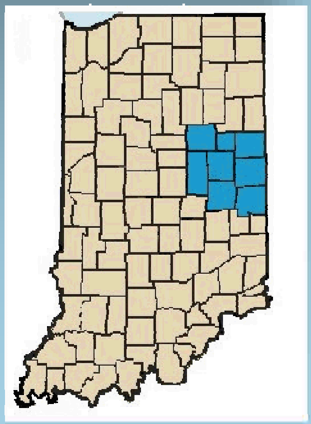
Presentation
American Society for Quality
Section 0904 - East Central Indiana

# Overview of Design of Experiments Principals

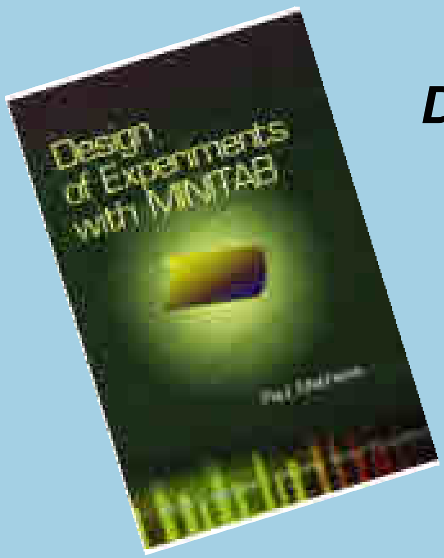

**Dr Rebecca Pierce**
**Associate Professor of Mathematical Sciences**
BALL STATE UNIVERSITY
EDUCATION REDEFINED

# Two-Way Experiments

Adapted from

***Design of Experiments with MINITAB***
by
Paul Mathews



# One Data Set - Same Analysis

- 4 Different Interpretations
- Which appropriate is determined by
  - Nature of study variables
  - Order data collected
- All 4 could be considered for 1 data set but only 1 of the 4 interpretations is correct


***Design of Experiments with MINITAB***

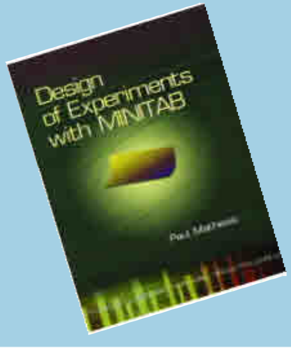
by

Paul Mathews

# Example

- Suppose we want to study manufacturing process where 5 different material lots are to be processed by 3 different operators
- Clearly a two-way experiment
- Whether we learn if there are differences between
  - Lots
  - Operators
  - Both
  - Nothing

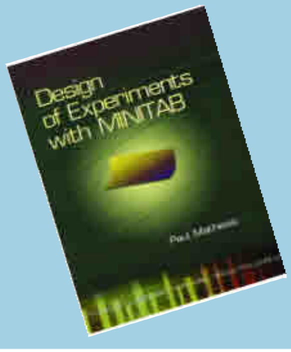
Determined by the order of the experimental runs




***Design of Experiments with MINITAB***
by
Paul Mathews

# Scenario 1

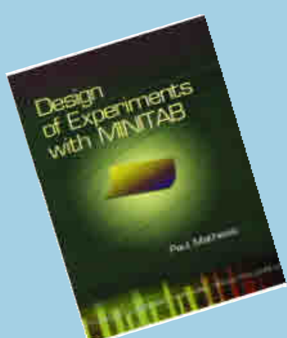
- If all 3 operators are on the job at same time & lots are processed in order (Lot 1, Lot 2,…,Lot 5)
- With each part being made by a randomly chosen operator, then
  - Lot is a blocking variable
  - Operator is the treatment variable
- Design - Randomized Block




***Design of Experiments with MINITAB***
by
Paul Mathews

# Scenario 2

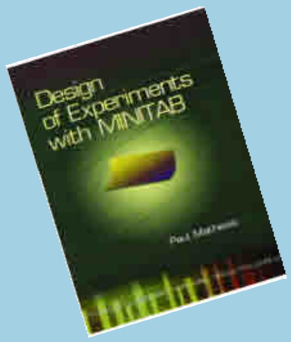
- Suppose the 3 operators work successive shifts & material from the 5 different lots can be processed in random order (Material is stockpiled & then chosen)
  - Lot is the treatment variable
  - Operator is a blocking variable
- Design - Randomized Block




***Design of Experiments with MINITAB***
by
Paul Mathews

# Scenario 3

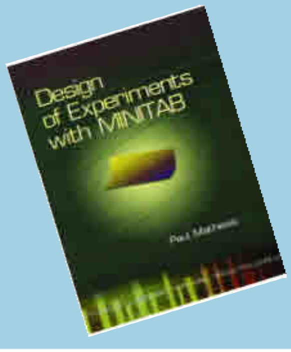
- Suppose the 3 operators & the 5 different lots are both completely randomized
- Possible if random parts from all 5 lots were given to random operators in random order
- Design – 5 x 3 Factorial
- Can identify differences between lots & operators in the same experiment




***Design of Experiments with MINITAB***
by
Paul Mathews

# Scenario 4

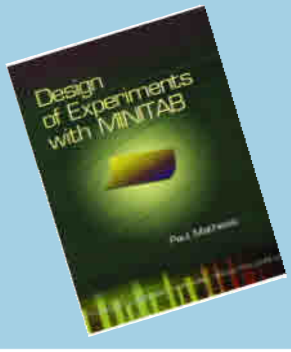
- Suppose neither the 3 operators nor the 5 different lots are randomized
- Easiest way to run the experiment
- But can't draw any safe conclusions about differences between level of lots or operators
- In effect both variables are blocking variables




***Design of Experiments with MINITAB***
by
Paul Mathews

# What happens?

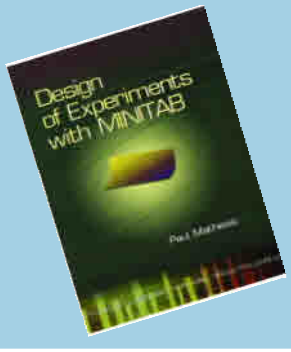
- For all 4 scenarios, data are exactly the same
- The ANOVA will give exactly the same results, since doesn’t account for the order of the experimental runs (random or not)
- Experimenter’s responsibility to see that the correct randomization plan is used
  - To support the goals of the experiment
  - To make the correct interpretation of the results




***Design of Experiments with MINITAB***
by
Paul Mathews

# Practice Example

- A two-way classification experiment is being designed
- The potential run orders being considered are shown in the next table
- Identify the design & describe what can be learned




***Design of Experiments with MINITAB***
by
Paul Mathews

# Two-way Experiment

| | | Run Order | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Plan** | **Variable** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
| 1 | A | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 |
| | B | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| 2 | A | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 |
| | B | 1 | 4 | 3 | 2 | 2 | 3 | 1 | 4 | 3 | 2 | 4 | 1 |
| 3 | A | 2 | 3 | 1 | 1 | 3 | 2 | 2 | 1 | 3 | 1 | 2 | 3 |
| | B | 1 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 3 | 4 | 4 | 4 |
| 4 | A | 2 | 3 | 2 | 1 | 1 | 2 | 3 | 1 | 3 | 2 | 1 | 3 |
| | B | 2 | 4 | 3 | 1 | 2 | 1 | 1 | 4 | 2 | 4 | 3 | 2 |

# Two-way Experiment

| | | | | | | | | | | | | | Run Order |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Plan** | **Variable** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
| 1 | A | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 |
| | B | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |

# Two-way Experiment

| | | Run Order | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Plan** | **Variable** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
| 1 | A | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 |
| | B | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| **2** | **A** | **1** | **1** | **1** | **1** | **2** | **2** | **2** | **2** | **3** | **3** | **3** | **3** |
| | **B** | **1** | **4** | **3** | **2** | **2** | **3** | **1** | **4** | **3** | **2** | **4** | **1** |

# Two-way Experiment

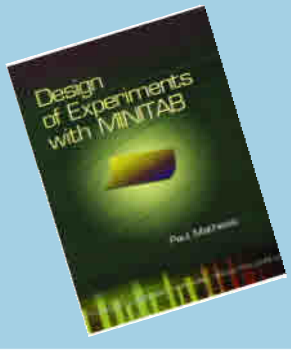
| | | Run Order | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Plan** | **Variable** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
| 1 | A | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 |
| | B | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| 2 | A | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 |
| | B | 1 | 4 | 3 | 2 | 2 | 3 | 1 | 4 | 3 | 2 | 4 | 1 |
| 3 | A | 2 | 3 | 1 | 1 | 3 | 2 | 2 | 1 | 3 | 1 | 2 | 3 |
| | B | 1 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 3 | 4 | 4 | 4 |

# Two-way Experiment

| | | Run Order | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Plan** | **Variable** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** |
| 1 | A | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 |
| | B | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| 2 | A | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 |
| | B | 1 | 4 | 3 | 2 | 2 | 3 | 1 | 4 | 3 | 2 | 4 | 1 |
| 3 | A | 2 | 3 | 1 | 1 | 3 | 2 | 2 | 1 | 3 | 1 | 2 | 3 |
| | B | 1 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 3 | 4 | 4 | 4 |
| 4 | A | 2 | 3 | 2 | 1 | 1 | 2 | 3 | 1 | 3 | 2 | 1 | 3 |
| | B | 2 | 4 | 3 | 1 | 2 | 1 | 1 | 4 | 2 | 4 | 3 | 2 |

# Factorial Designs

- In general, can be used to analyze multi-way classification problems
- There will be two-way interactions between pairs of variables
- Potentially there will be three-way & higher order interactions to consider

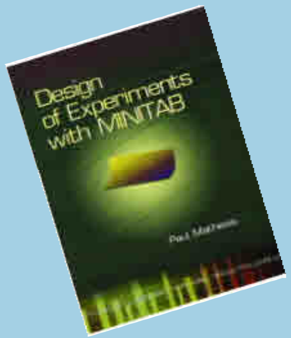



***Design of Experiments with MINITAB***
by
Paul Mathews

# Factorial Designs

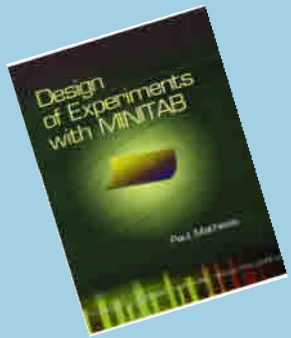
- Number of classification variables is k
  - # of Main Effects $\binom{k}{1}$
  - # of 2-way Interactions $\binom{k}{2}$
  - # of 3-way interactions $\binom{k}{3}$
  - and so on


***Design of Experiments with MINITAB***
by
Paul Mathews




# Factorial Designs

- Number of degrees of freedom for any interaction is always equal to the product of the # of df of the main effects involved in the interaction
- Most engineering situations 3-way & higher interactions are rare
  - Usually assume insignificant
  - Always get opinion of expert on the process

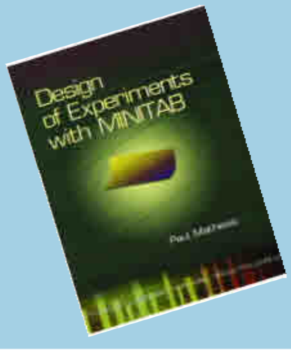



# Factorial Designs

- If 2 or more replicates, then all interactions between 2 variables, 3 variables, & up to the single k-factor interaction can be included in the model
- Wise to analyze the full model to see if any higher-order interactions are significant
  - If not, drop them from model & pool with error
  - Simpler to explain

# Factorial Designs

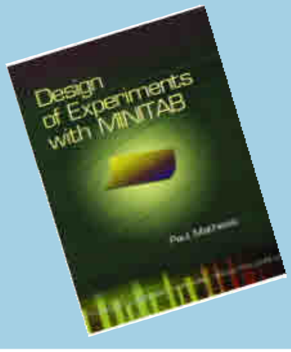
- Same # of replicates for each cell is said to be balanced
- When all possible a x b x … cells are included in the design, factorial design is said to be full
- These among the best-behaved designs
- Some of the easiest to analyze




***Design of Experiments with MINITAB***
by
Paul Mathews

# 3 x 4 x 5 Full Factorial Experiment

- With two replicates
- Determine the df for each term in model if the model includes main effects as well as 2-factor and 3-factor interactions.
- How many error df will there be?
- How many error df if we omit the 3-factor interaction from the model?




***Design of Experiments with MINITAB***
by
Paul Mathews

# 3 x 4 x 5 Full Factorial Experiment with 2 replicates

| **Source** | **df** | |
|---|---|---|
| A | 2 | 2 |
| B | 3 | 3 |
| C | 4 | 4 |
| AB | 6 | 6 |
| AC | 8 | 8 |
| BC | 12 | 12 |
| ABC | 24 | |
| $\epsilon$ | 60 | 84 |
| Total | 119 | 119 |

# Introduction to Factorials

- **General principles** of factorial experiments
- The **two-factor factorial** with fixed effects
- The **ANOVA** for factorials
- Extensions to more than two factors
- **Quantitative** and **qualitative** factors – response curves and surfaces

# Some Basic Definitions

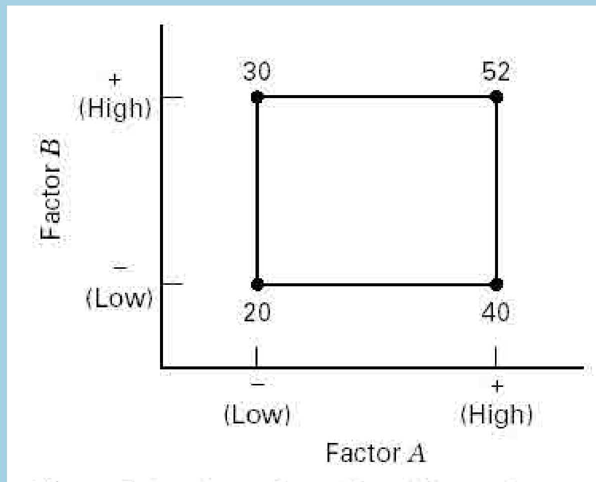



**Figure 5-1** A two-factor factorial experiment, with the response (y) shown at the corners.

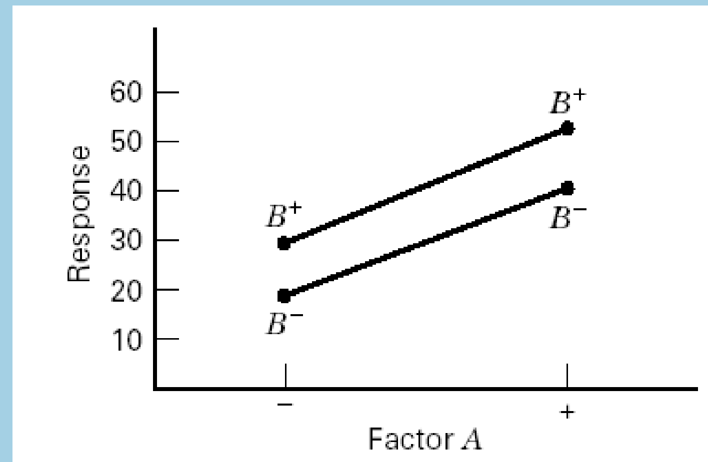



**Figure 5-3** A factorial experiment without interaction.

**Definition of a factor effect: The change in the mean response when the factor is changed from low to high**

$$A = \bar{y}_{A^+} - \bar{y}_{A^-} = \frac{40+52}{2} - \frac{20+30}{2} = 21$$

$$B = \bar{y}_{B^+} - \bar{y}_{B^-} = \frac{30+52}{2} - \frac{20+40}{2} = 11$$

$$AB = \frac{52+20}{2} - \frac{30+40}{2} = -1$$

# The Case of Interaction:

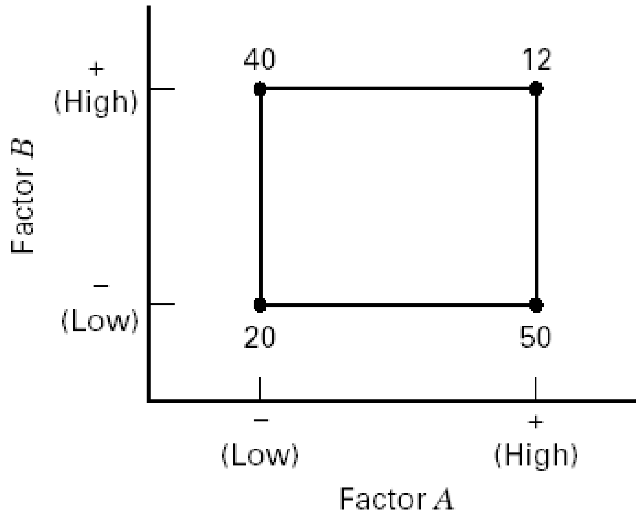



**Figure 5-2** A two-factor factorial experiment with interaction.

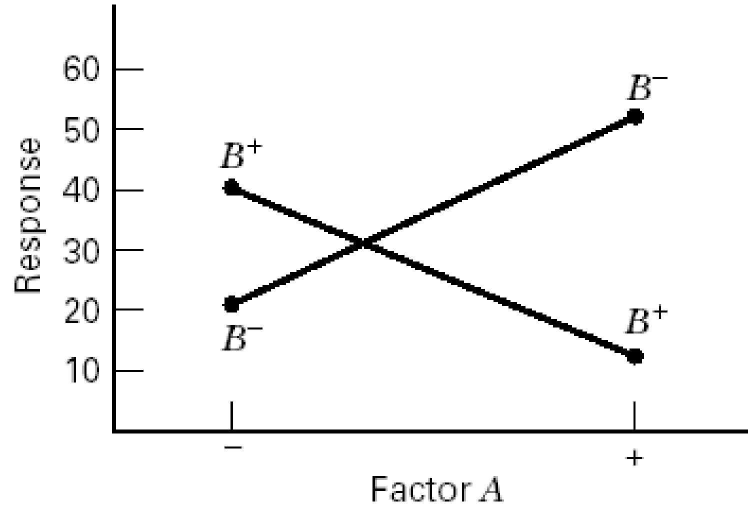



**Figure 5-4** A factorial experiment with interaction.

$$A = \bar{y}_{A^+} - \bar{y}_{A^-} = \frac{50+12}{2} - \frac{20+40}{2} = 1$$

$$B = \bar{y}_{B^+} - \bar{y}_{B^-} = \frac{40+12}{2} - \frac{20+50}{2} = -9$$

$$AB = \frac{12+20}{2} - \frac{40+50}{2} = -29$$

# Regression Model & Associated Response Surface

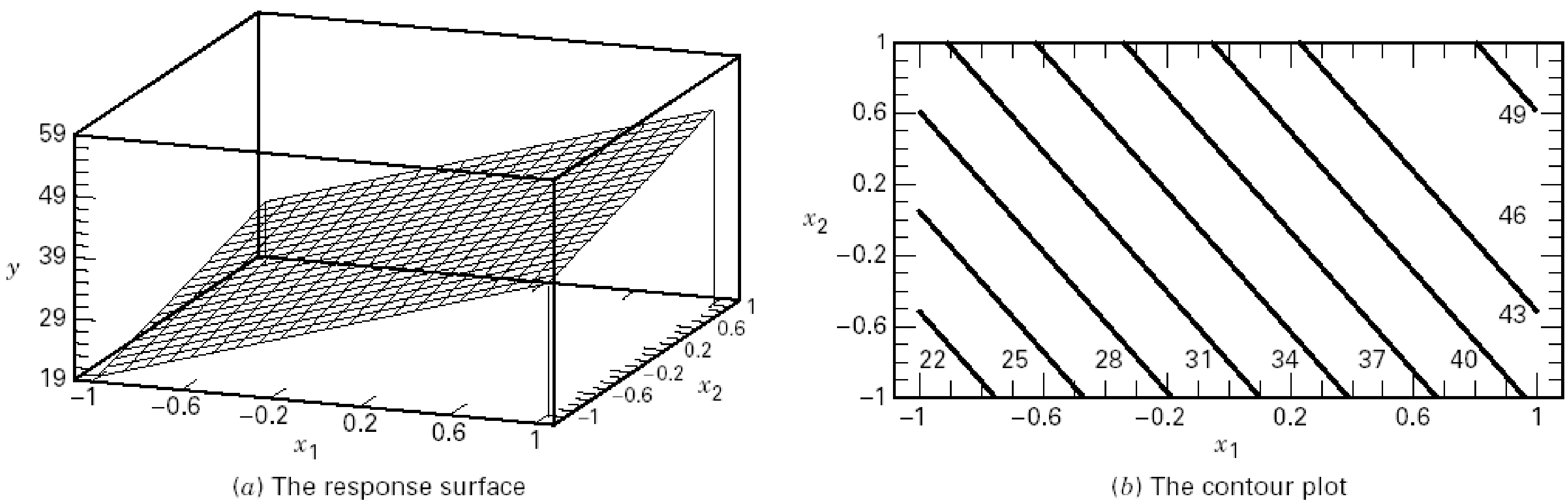



**Figure 5-5** Response surface and contour plot for the model $\hat{y} = 35.5 + 10.5x_1 + 5.5x_2$.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$$

The least squares fit is

$$\hat{y} = 35.5 + 10.5x_1 + 5.5x_2 + 0.5x_1x_2$$

$$\cong 35.5 + 10.5x_1 + 5.5x_2$$

# Effect of Interaction on Response Surface

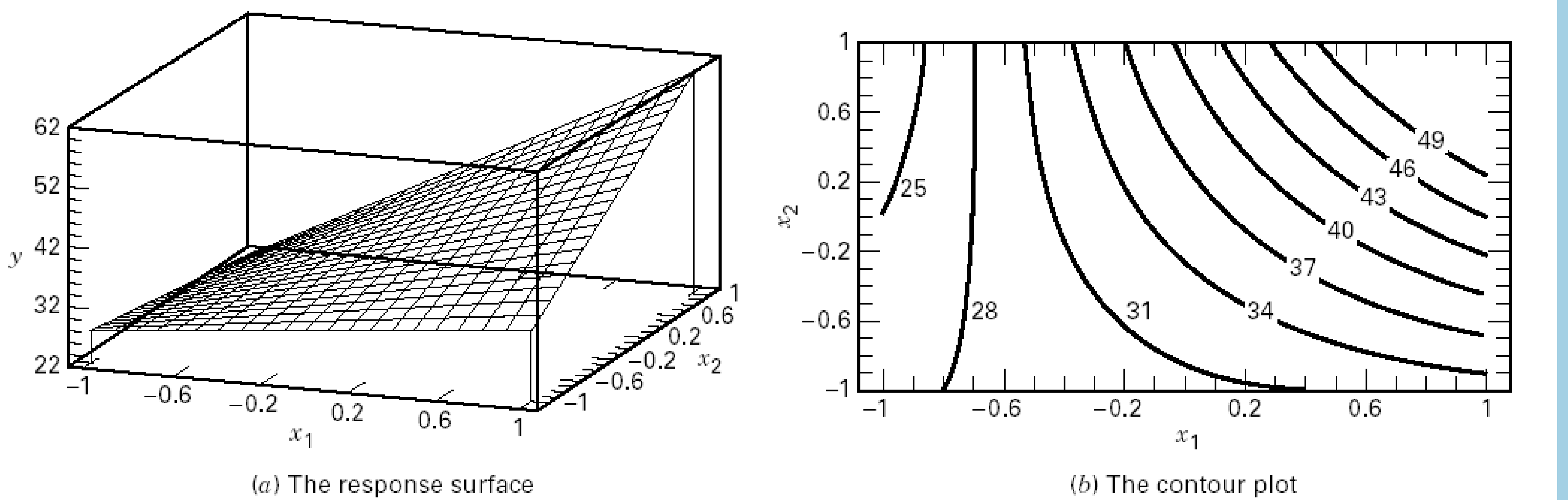



**Figure 5-6** Response surface and contour plot for the model $\hat{y} = 35.5 + 10.5x_1 + 5.5x_2 + 8x_1x_2$.

Suppose that we add an interaction term to the model:

$$\hat{y} = 35.5 + 10.5x_1 + 5.5x_2 + 8x_1x_2$$

**Interaction** is actually a form of **curvature**

# Advantages

- More Efficient than one-factor-at-a-time experiments
- Necessary when interaction may be present to avoid misleading conclusions
- Allow the effects of a factor to be estimated at several levels of the other factor, yielding conclusions that are valid over a range of experimental conditions

# Example 5-1
# The Battery Life Experiment

Table 5-1 Life (in hours) Data for the Battery Design Example

| Material Type | Temperature (°F) 15 | | 70 | | 125 | |
|---|---|---|---|---|---|---|
| 1 | 130 | 155 | 34 | 40 | 20 | 70 |
| | 74 | 180 | 80 | 75 | 82 | 58 |
| 2 | 150 | 188 | 136 | 122 | 25 | 70 |
| | 159 | 126 | 106 | 115 | 58 | 45 |
| 3 | 138 | 110 | 174 | 120 | 96 | 104 |
| | 168 | 160 | 150 | 139 | 82 | 60 |

*A* = Material type; *B* = Temperature (A **quantitative** variable)

1. What **effects** do material type & temperature have on life?
2. Is there a choice of material that would give long life ***regardless of temperature*** (a **robust** product)?

# General Two-Factor Factorial Experiment

**Table 5-2** General Arrangement for a Two-Factor Factorial Design

| | | Factor *B* — *b* levels | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | . . . | *b* |
| Factor *A* | 1 | $y_{111}, y_{112}, \ldots, y_{11n}$ | $y_{121}, y_{122}, \ldots, y_{12n}$ | | $y_{1b1}, y_{1b2}, \ldots, y_{1bn}$ |
| | 2 | $y_{211}, y_{212}, \ldots, y_{21n}$ | $y_{221}, y_{222}, \ldots, y_{22n}$ | | $y_{2b1}, y_{2b2}, \ldots, y_{2bn}$ |
| *a* levels | ⋮ | | *n* replicates | | |
| | *a* | $y_{a11}, y_{a12}, \ldots, y_{a1n}$ | $y_{a21}, y_{a22}, \ldots, y_{a2n}$ | | $y_{ab1}, y_{ab2}, \ldots, y_{abn}$ |

This is a **completely randomized design**

Table 5-2 General Arrangement for a Two-Factor Factorial Design

| | | Factor B | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | ... | $b$ |
| Factor A | 1 | $y_{111}, y_{112}, \ldots, y_{11n}$ | $y_{121}, y_{122}, \ldots, y_{12n}$ | | $y_{1b1}, y_{1b2}, \ldots, y_{1bn}$ |
| | 2 | $y_{211}, y_{212}, \ldots, y_{21n}$ | $y_{221}, y_{222}, \ldots, y_{22n}$ | | $y_{2b1}, y_{2b2}, \ldots, y_{2bn}$ |
| | ⋮ | | | | |
| | $a$ | $y_{a11}, y_{a12}, \ldots, y_{a1n}$ | $y_{a21}, y_{a22}, \ldots, y_{a2n}$ | | $y_{ab1}, y_{ab2}, \ldots, y_{abn}$ |

Statistical (effects) model:

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \\ k = 1, 2, \ldots, n \end{cases}$$

# Extension of the ANOVA to Factorials (Fixed Effects Case)

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{...})^2 = bn\sum_{i=1}^{a}(\bar{y}_{i..}-\bar{y}_{...})^2 + an\sum_{j=1}^{b}(\bar{y}_{.j.}-\bar{y}_{...})^2$$

$$+n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}+\bar{y}_{...})^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{ij.})^2$$

$$SS_T = SS_A + SS_B + SS_{AB} + SS_E$$

$df$ breakdown:

$$abn-1 = a-1+b-1+(a-1)(b-1)+ab(n-1)$$

# ANOVA Table – Fixed Effects Case

Table 5-3 The Analysis of Variance Table for the Two-Factor Factorial, Fixed Effects Model

| Source of Variation | Sum of Squares | Degrees of Freedom | Mean Square | $F_0$ |
|---|---|---|---|---|
| $A$ treatments | $SS_A$ | $a-1$ | $MS_A = \frac{SS_A}{a-1}$ | $F_0 = \frac{MS_A}{MS_E}$ |
| $B$ treatments | $SS_B$ | $b-1$ | $MS_B = \frac{SS_B}{b-1}$ | $F_0 = \frac{MS_B}{MS_E}$ |
| Interaction | $SS_{AB}$ | $(a-1)(b-1)$ | $MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$ | $F_0 = \frac{MS_{AB}}{MS_E}$ |
| Error | $SS_E$ | $ab(n-1)$ | $MS_E = \frac{SS_E}{ab(n-1)}$ | |
| Total | $SS_T$ | $abn-1$ | | |

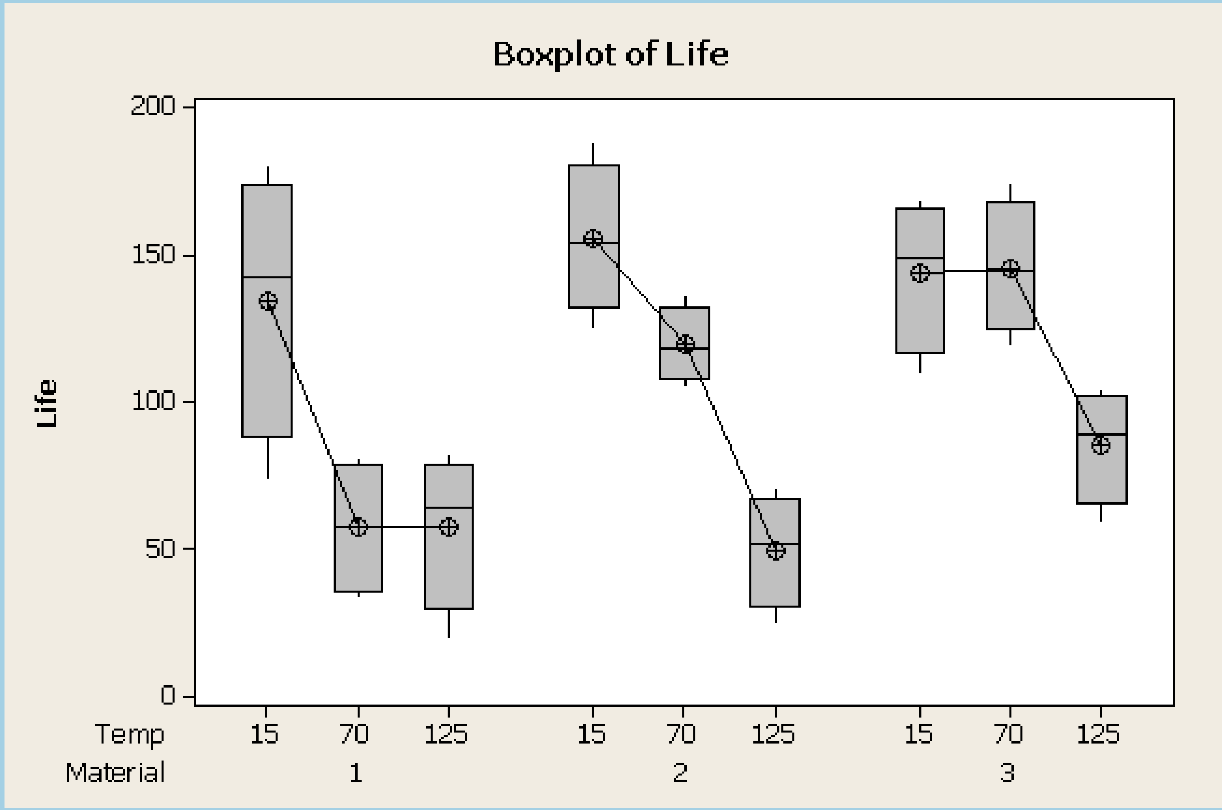
**Software** will perform the computations

Details of **manual computing** – see pp. 169 & 170

# Example 5-1

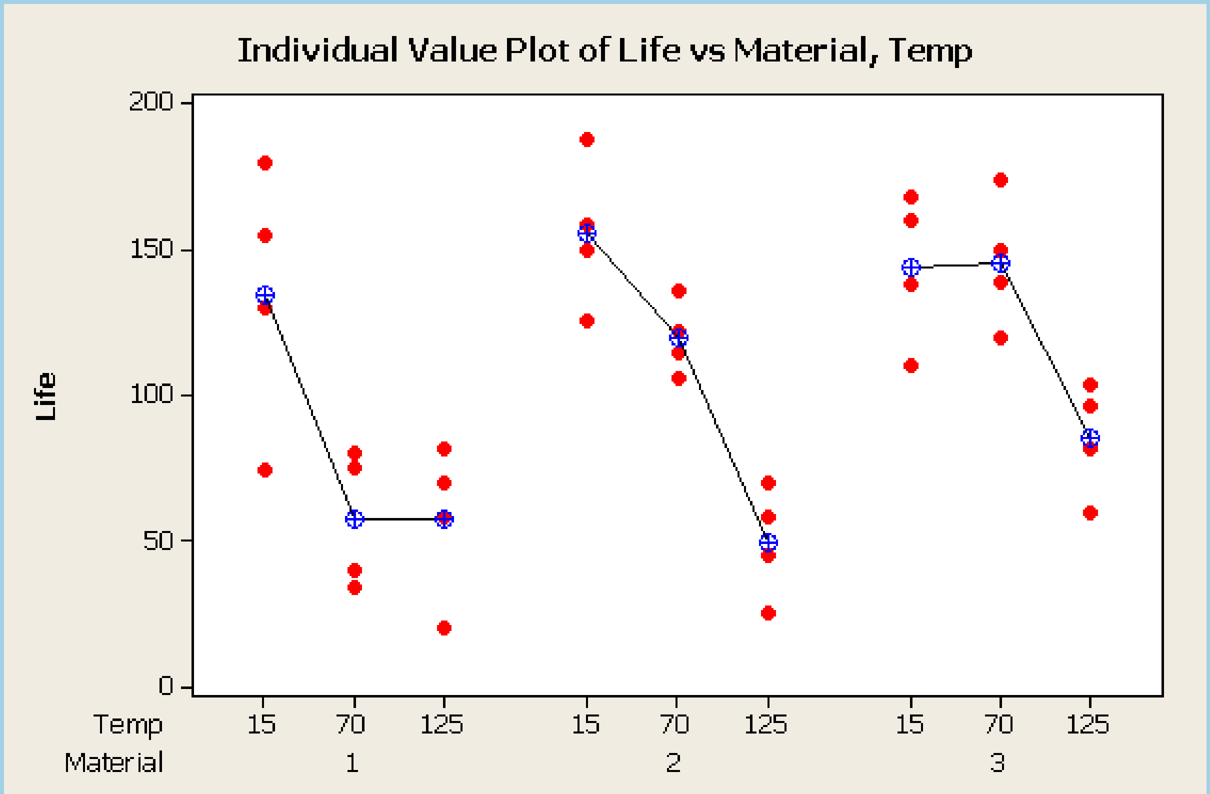
Worksheet 2 ***

| ↓ | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 |
|---|---|---|---|---|---|---|---|---|
| | Material | Temp | Life | | | | | |
| 1 | 1 | 15 | 130 | | | | | |
| 2 | 1 | 15 | 155 | | | | | |
| 3 | 1 | 15 | 74 | | | | | |
| 4 | 1 | 15 | 180 | | | | | |
| 5 | 1 | 70 | 34 | | | | | |
| 6 | 1 | 70 | 40 | | | | | |
| 7 | 1 | 70 | 80 | | | | | |
| 8 | 1 | 70 | 75 | | | | | |
| 9 | 1 | 125 | 20 | | | | | |
| 10 | 1 | 125 | 70 | | | | | |
| 11 | 1 | 125 | 82 | | | | | |
| 12 | 1 | 125 | 58 | | | | | |
| 13 | 2 | 15 | 150 | | | | | |
| 14 | 2 | 15 | 188 | | | | | |
| 15 | 2 | 15 | 159 | | | | | |
| 16 | 2 | 15 | 126 | | | | | |




**Boxplot**

A graphical summary of the distribution of a sample that shows its shape, central tendency, and variability.




**Individual value plot**

Use to assess and compare sample distributions through individual data values, with optional grouping by categorical variables. An individual value plot can also help you to check for obvious mistakes in coding values

The individual value plot is similar to a boxplot in (informally) identifying outliers and distribution shape, but it is unique in that it graphs each value separately. This is especially useful when you have relatively few observations or when it's important to evaluate the effect of each observation.

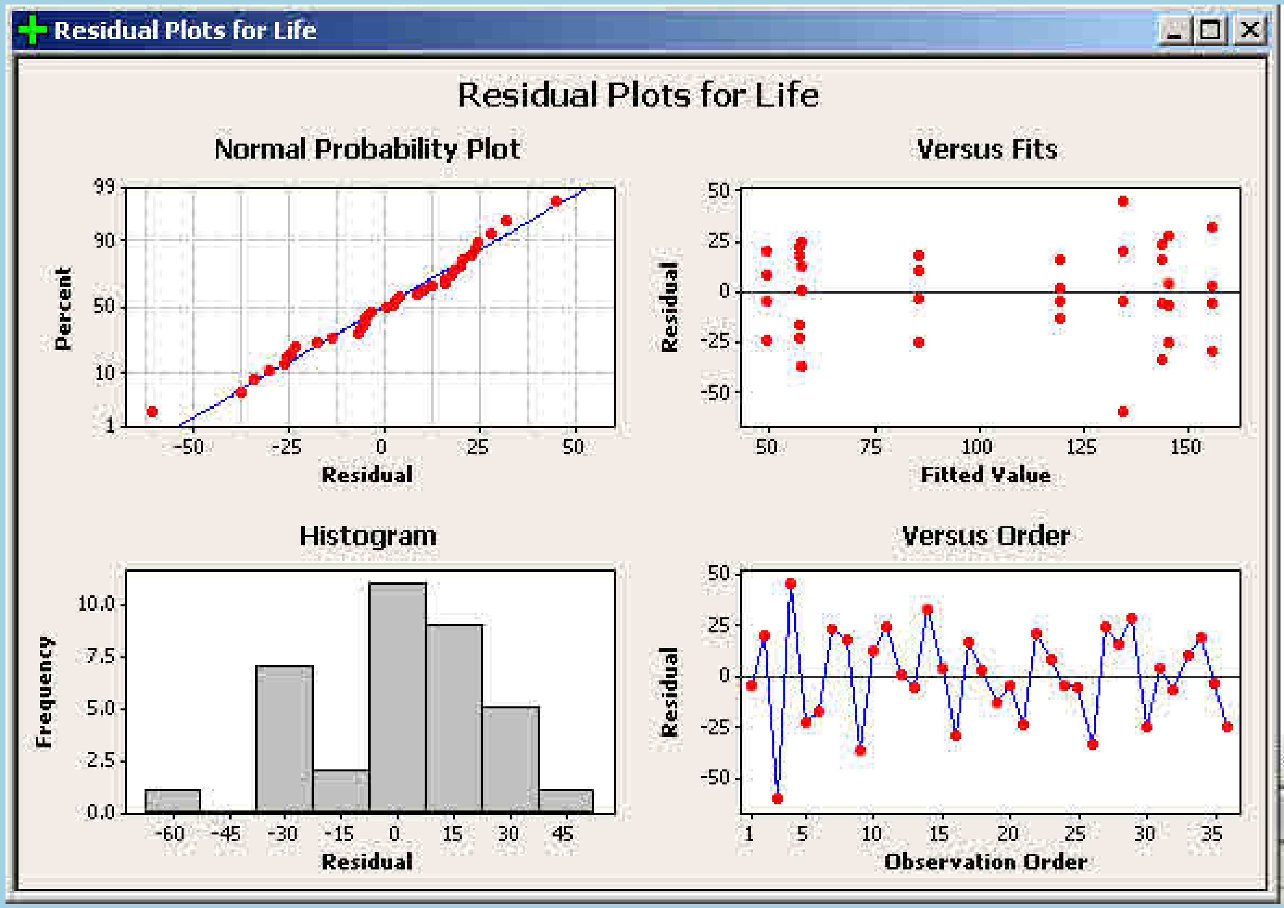
# MINITAB Output – Example 5-1

**Two-way ANOVA: Life versus Material, Temp**

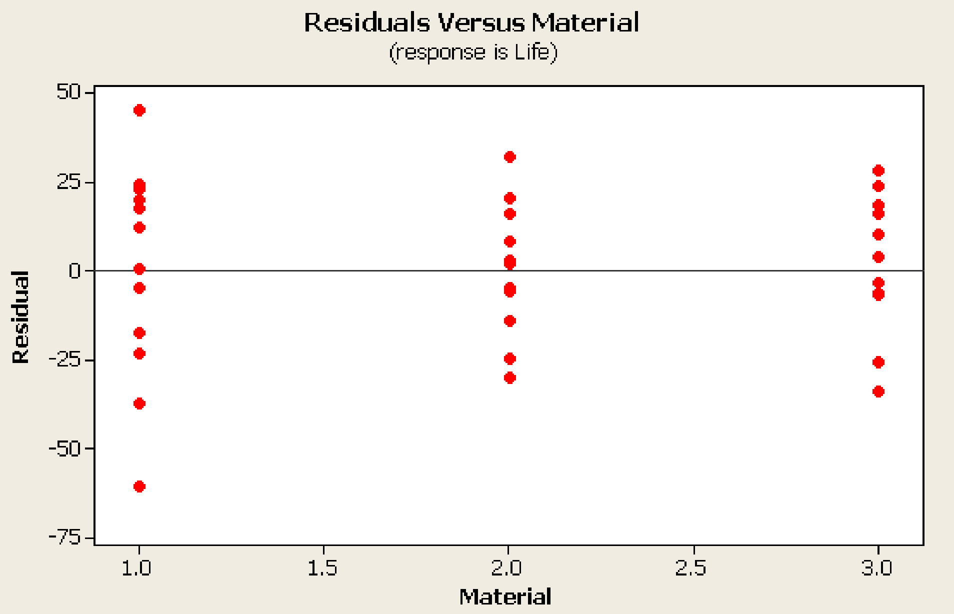
| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Material | 2 | 10683.7 | 5341.9 | 7.91 | 0.002 |
| Temp | 2 | 39118.7 | 19559.4 | 28.97 | 0.000 |
| Interaction | 4 | 9613.8 | 2403.4 | 3.56 | 0.019 |
| Error | 27 | 18230.8 | 675.2 | | |
| Total | 35 | 77647.0 | | | |

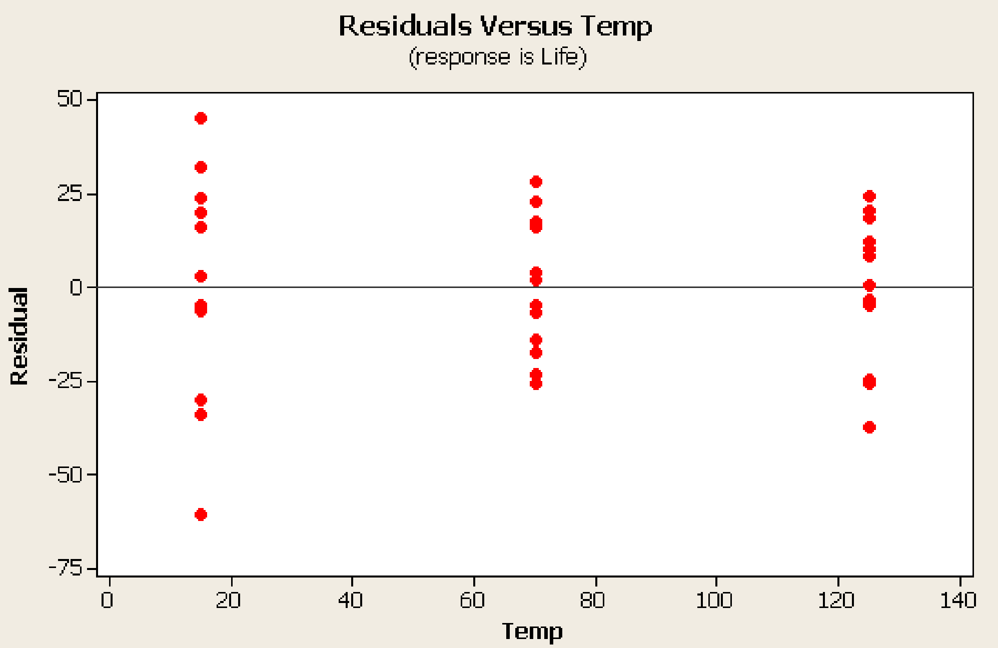
S = 25.98   R-Sq = 76.52%   R-Sq(adj) = 69.56%

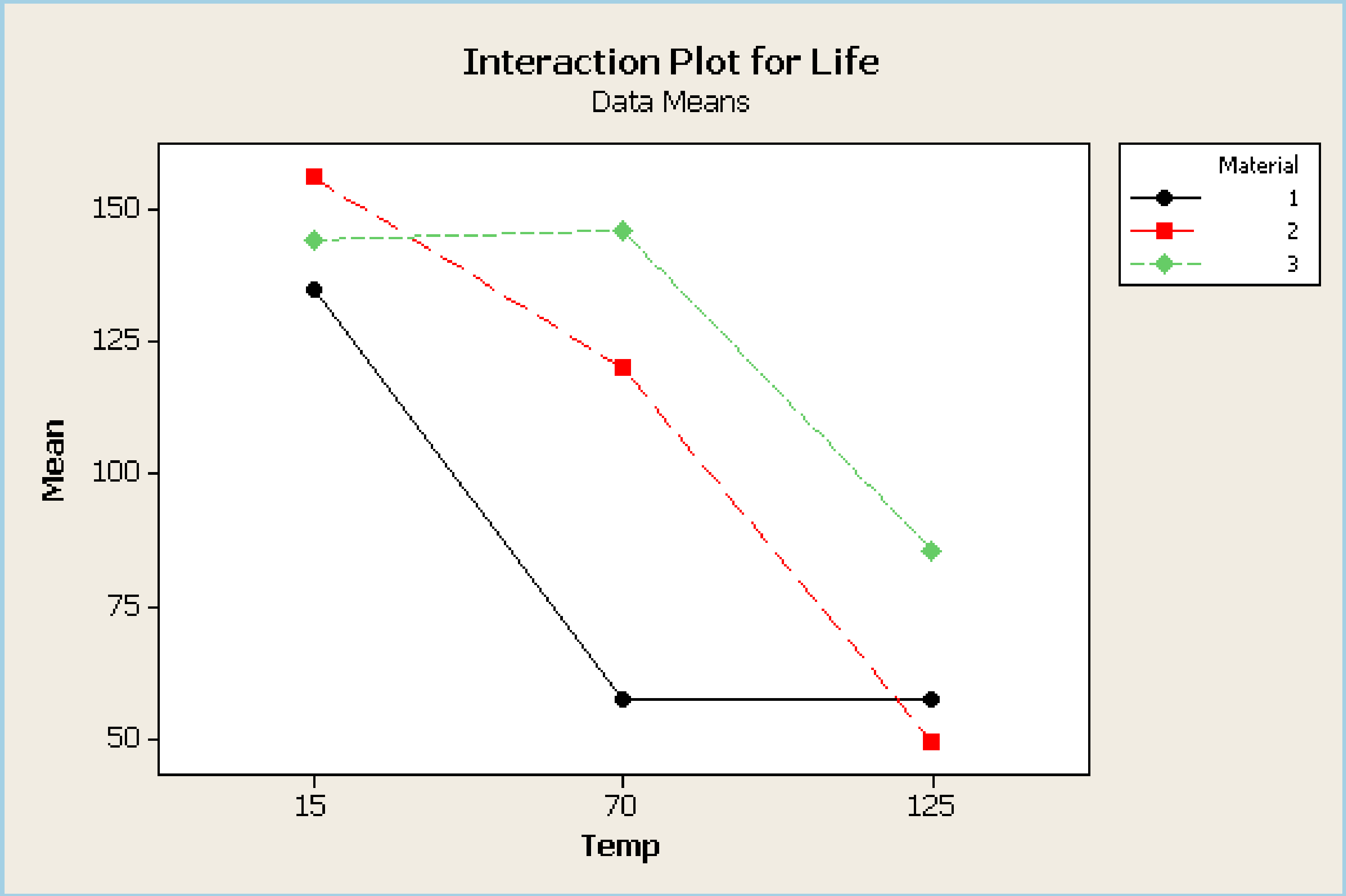
# Residual Analysis – Example 5-1



# Residual Analysis – Example 5-1

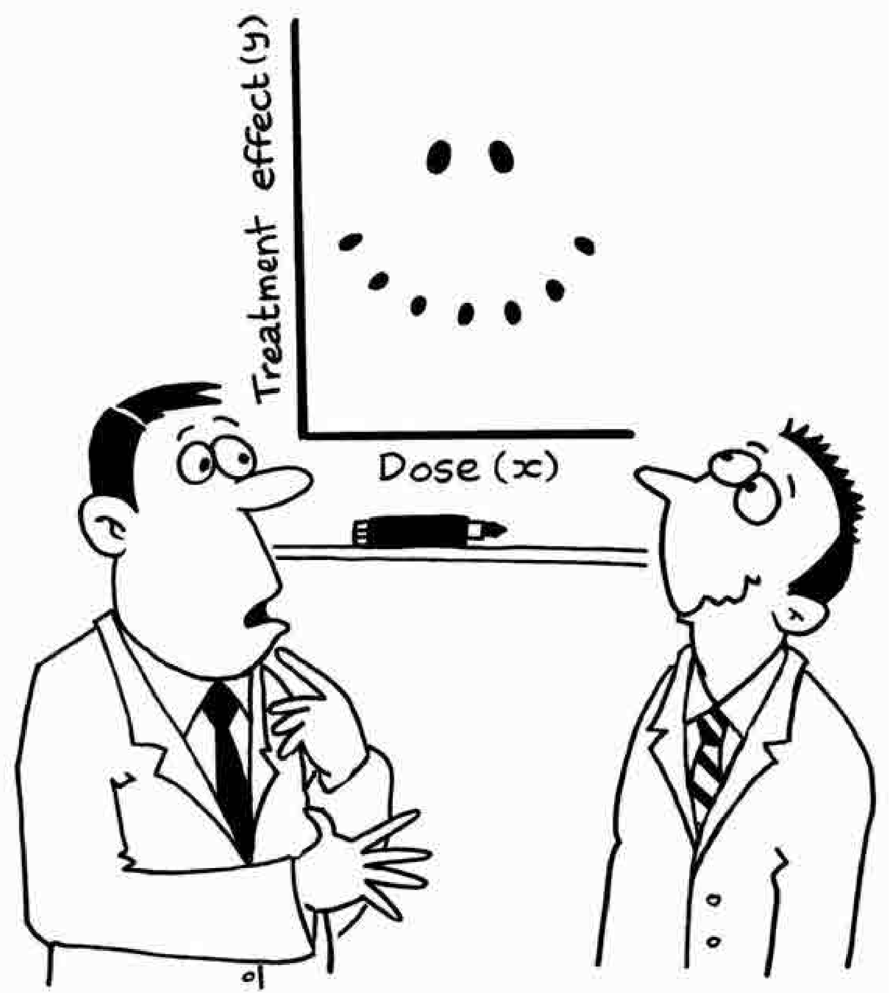





Interactions Plot creates a single interaction plot for two factors, or a matrix of interaction plots for three to nine factors. An interactions plot is a plot of means for each level of a factor with the level of a second factor held constant. Interactions plots are useful for judging the presence of interaction.

Interaction is present when the response at a factor level depends upon the level(s) of other factors. Parallel lines in an interactions plot indicate no interaction. The greater the departure of the lines from the parallel state, the higher the degree of interaction. To use interactions plot, data must be available from all combinations of levels.




"It's a non-linear pattern with outliers.....but for some reason I'm very happy with the data."