A Potpourri of Advanced Topics in Six Sigma

By Thomas Pyzdek

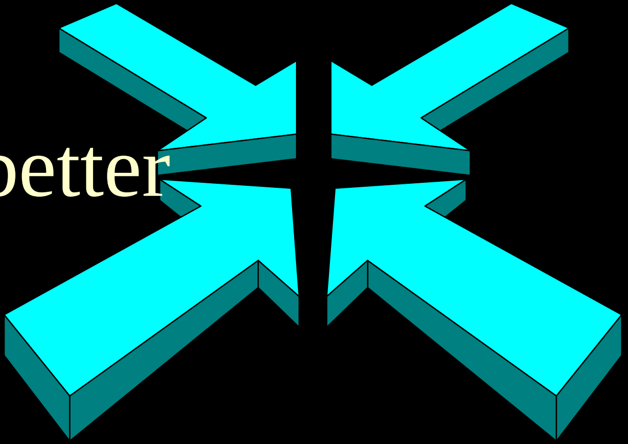
ASQ Six Sigma Conference

January 22-23, 2001

San Diego, CA

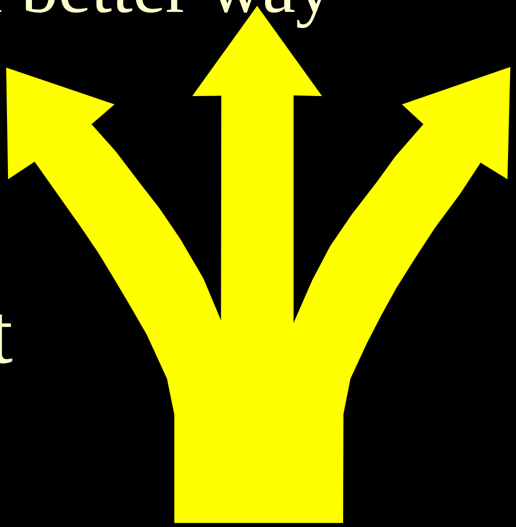
Anticipation

- Control by a central mind
- Predict and prevent potential dangers before damage is done
- Preserve stability
- The less fluctuation, the better
- Goal: prevent failures
- Bug-chasing at BayBank



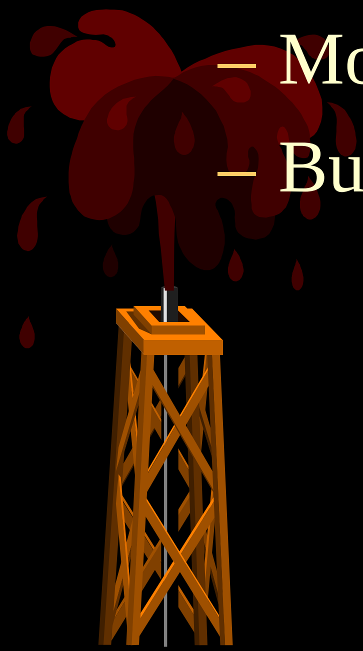
Resilience

- Experimentation
  - Try a lot of things and keep what works
  - Premise: there is always a better way
  - Goal: many failures
- The edge of chaos
- Interleaf versus Microsoft



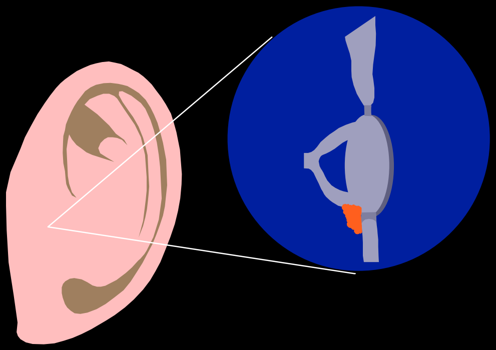
Efficiency

- Hallmarks: order and simplicity
  – Model of efficiency: bacteria
  – Business example: sucker rod manufacturing



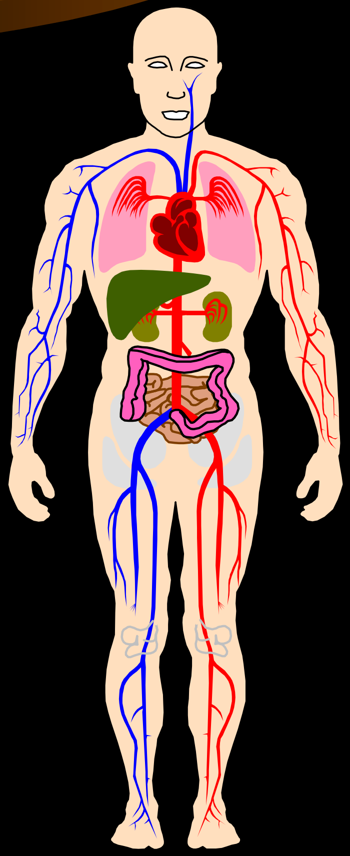
Creativity

- Hallmarks: messiness and complexity
- Sloppiness, poor fit, quirky design, redundancy
- Slack
  - Many great thinkers did their best work when they had time off
  - Ford's efficiency expert
  - Why Deming loved monopolies



Complex Adaptive Systems (CAS)

- Order without control
- Interacting agents described by rules
  - Firms, species, neurons
- Order and complexity emerge
  - Free economies
  - Ecosystems
  - The brain



The Control Paradox

- The fatal conceit
- Who plans America?
  - No one
  - Everyone
- Nucor Vs. Big Steel's HQ
- Control discourages adaptation

Adaptive agents

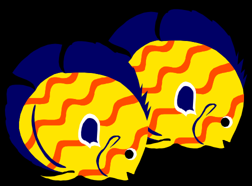
- Defined by performance rules
  - Detectors: filter information from their environment

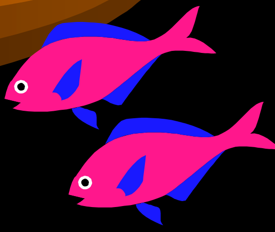
  - Effectors: actions taken in response to information processing

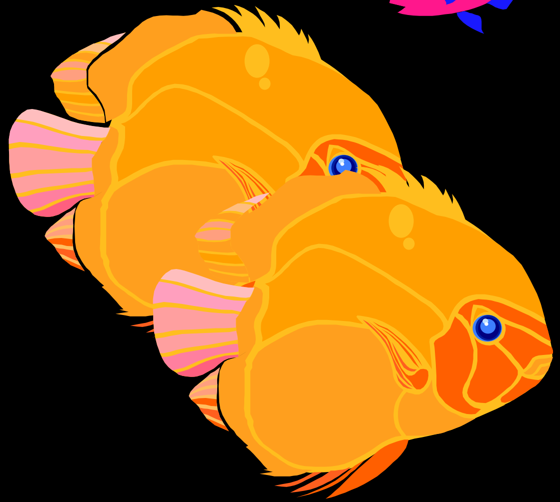
- Adapt by discovering new rules

Mechanisms of CAS

- Tagging
- Internal models
  - Tacit
  - Overt
- Building blocks

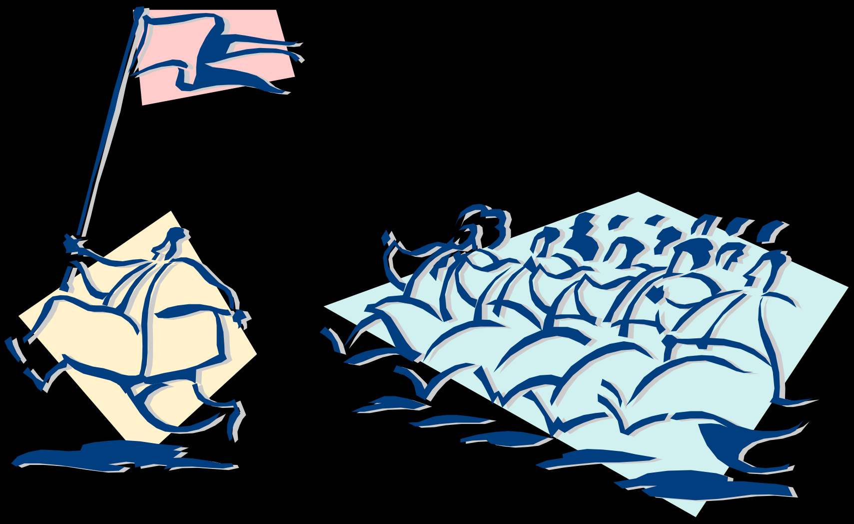






Properties of CAS-1

- Aggregation
- Nonlinearity
  - Multiplier effect
- Flows



Properties of CAS-2

- Diversity
  - Convergence
  - Mimicry

Genetic Learning-1

- Knowledge encoded in a schema of rules
- Rules modified by experience (learning)
  - Six sigma black belt helpful, but not required!
- Knowledge shared through intercourse
  - Sexual or social sharing
  - Crossover
  - Mutation

Genetic Learning-2

- The Prisoner's Dilemma Game

		Agent 1	
		Cooperate	Defect
Agent 2	Cooperate	\$3, \$3	\$0, \$5
	Defect	\$5, \$0	\$1, \$1

Genetic Learning-3

- *Evolving* strategy versus planning strategy

c d c c d d c c

. . .

d d d c c d d c

Play 64 times, with 20 players, “mate” winners. Repeat 151 times.

---

c d c c d d c c

d d d c c d d c

c d d c d d c c

c d c c d d d d

Crossover and Mutation

“Mutant”

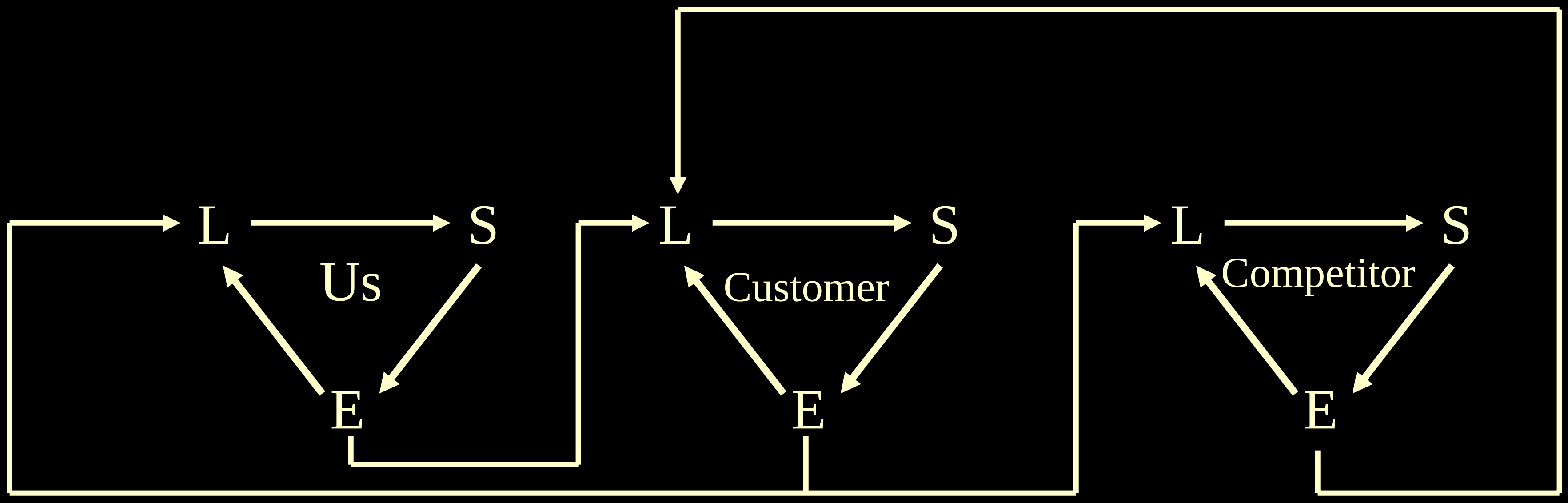
Cultivation Techniques

- Landscape Theory
  - Aggregation building within organizations
  - Groups highly compatible agents
  - Determines which configurations are stable
- Norm building techniques
  - Creating internal models among agents
  - Aids in tagging and aggregation

What Can Six Sigma Professionals Do?

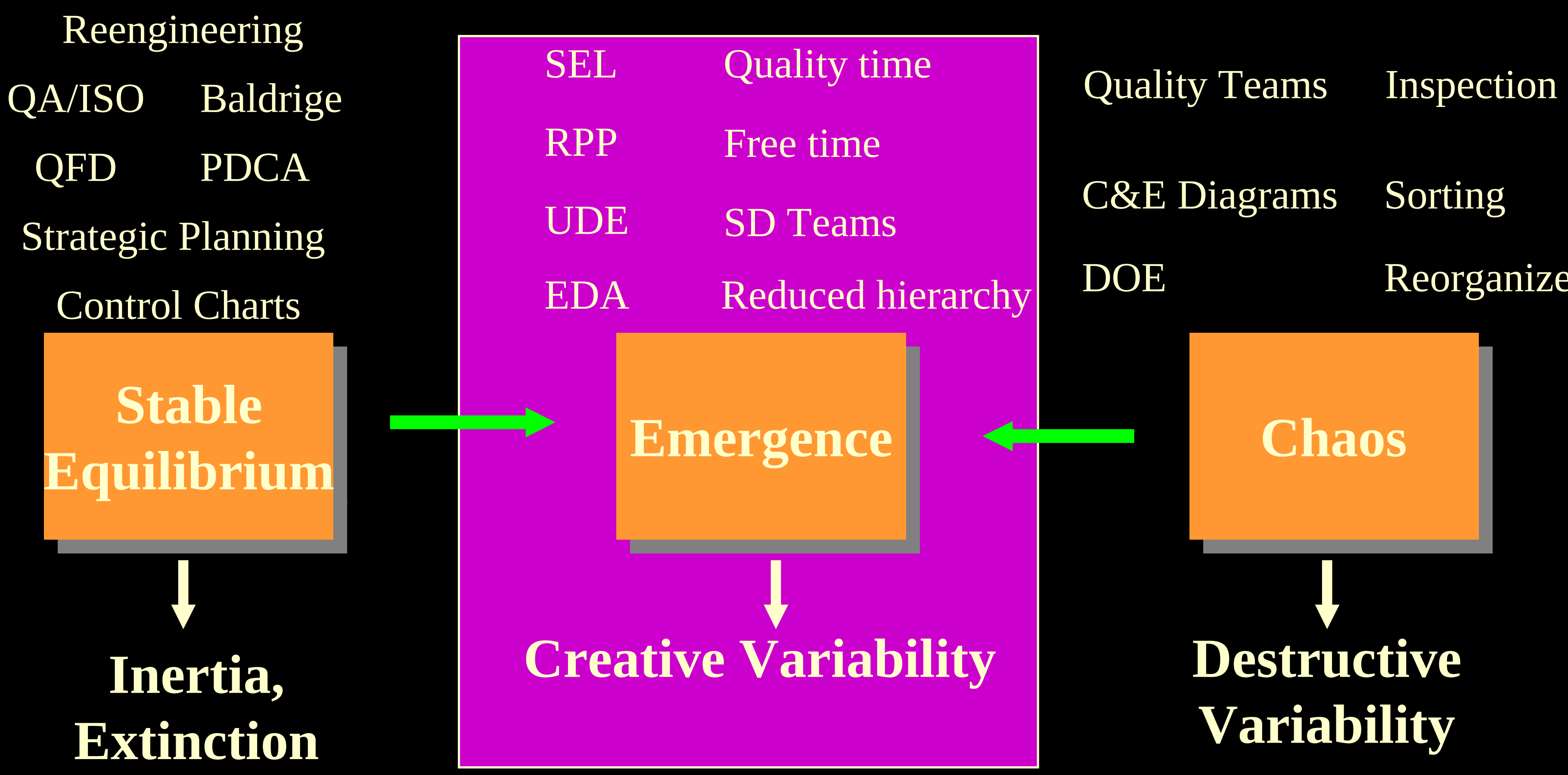
- Challenge assumptions, admit mistakes
- Develop resilient approaches to quality
- Tolerate variation, failure, redundancy, slop
- Value individuals as well as groups
- Use special knowledge to accelerate learning

SEL
A New Approach to Quality

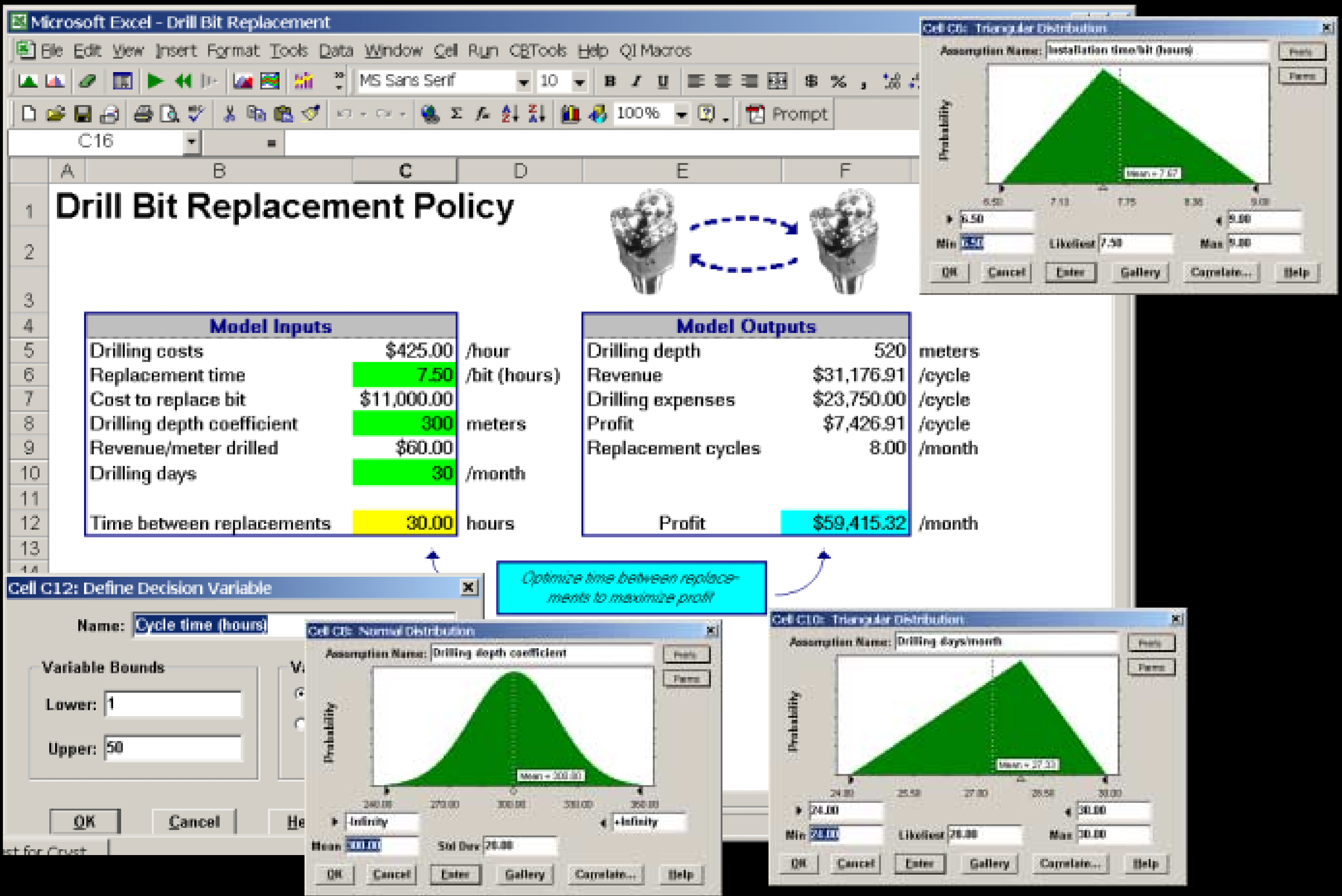


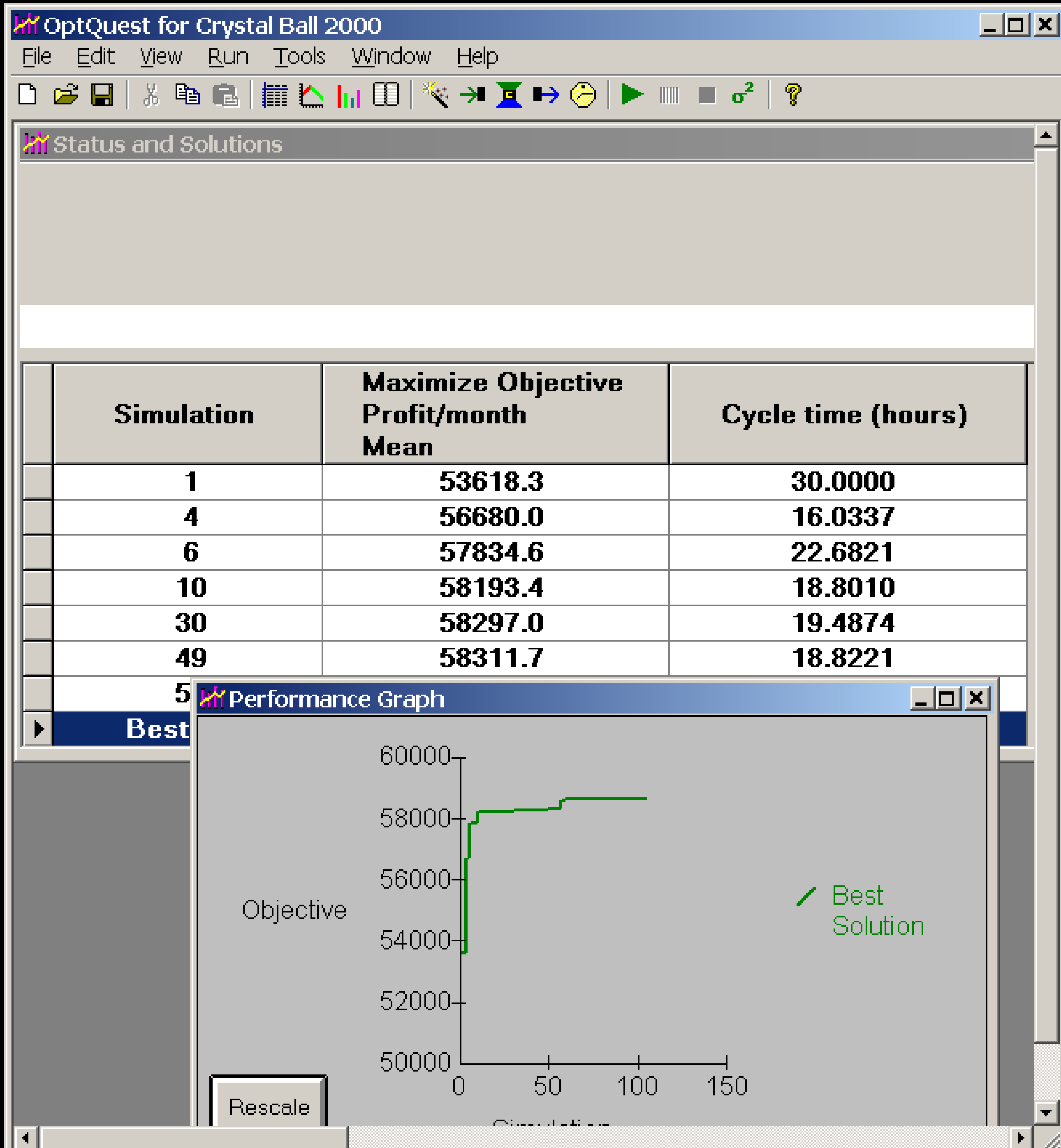
Dynamic SEL Model

The New Quality Paradigm



Practical Problem: Finding best replacement interval when inputs vary

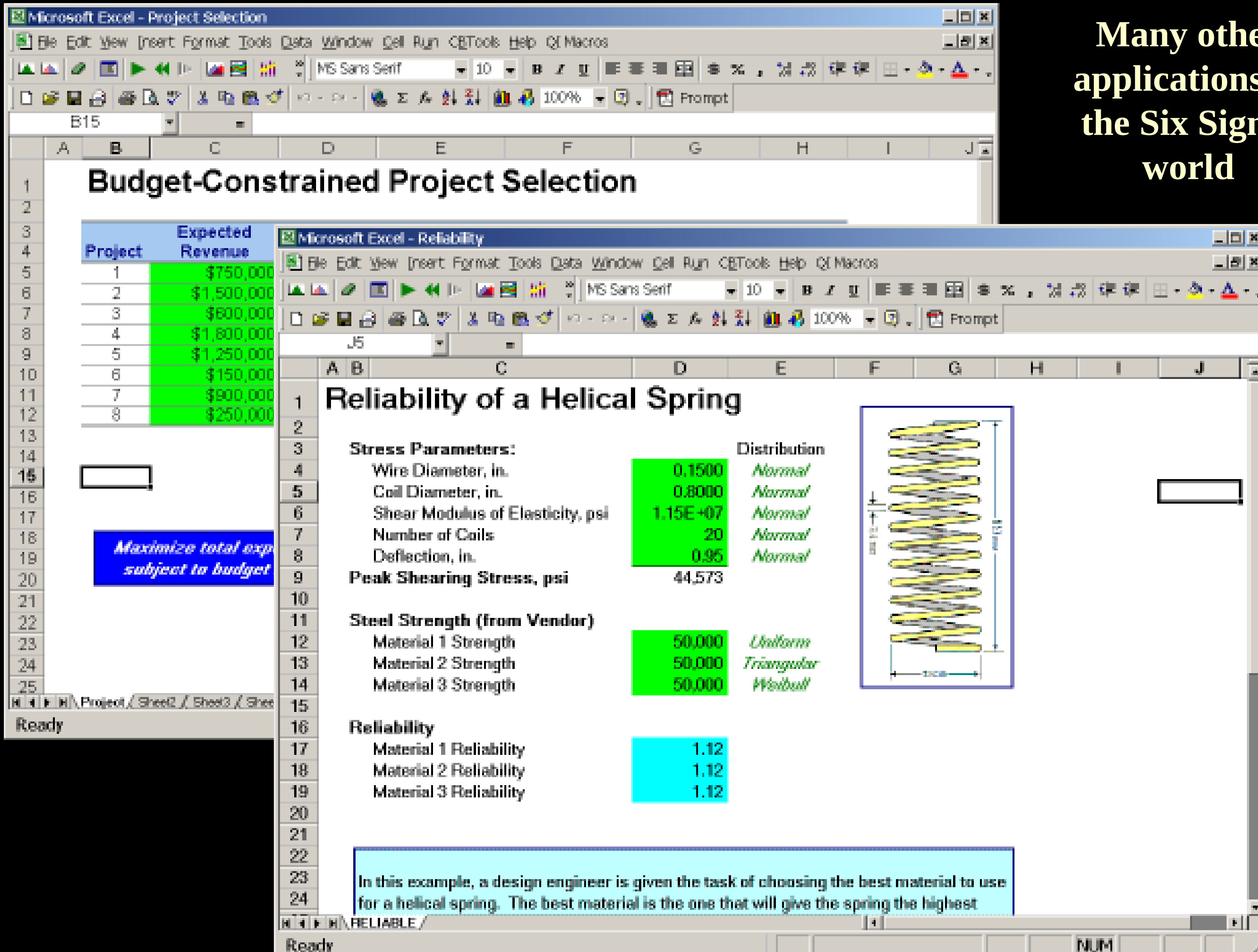




Simulation	Maximize Objective Profit/month Mean	Cycle time (hours)
1	53618.3	30.0000
4	56680.0	16.0337
6	57834.6	22.6821
10	58193.4	18.8010
30	58297.0	19.4874
49	58311.7	18.8221
5		
Best		

Solution found using genetic search algorithms to explore alternatives (simulated SEL)

Many other applications in the Six Sigma world



Data Mining for Quality

What is Data Mining?

The exploration and analysis, by automatic or semiautomatic means, of large quantities of data in order to discover meaningful patterns and rules.

Why Data Mining is Becoming Popular

- Vast quantities of data are being produced
- Data are being warehoused
- Computing power is affordable
- Competitive pressures
  - Service a key competitive factor
  - “Mass customization”
  - Information a product in itself

Data Mining Tasks

- Classification
- Estimation
- Prediction
- Affinity grouping
- Clustering
- Description

Data Mining Approaches

- Top-down *hypothesis testing*
  - Retrospective testing of preconceived ideas
- Bottom-up *knowledge discovery* (KD)
  - Directed: Explain or categorize a particular data field (e.g., income)
  - Undirected: Search for patterns or similarities among groups of records

Hypothesis Testing

- Generate good ideas
- Determine data needed to test hypotheses
- Locate data
- Prepare data for analysis
- Build computer models using data
- Use computer models to confirm/disconfirm hypotheses

Directed Knowledge Discovery-1

- Goal-directed
- Answer questions such as:
  - Which products have the highest failure rates?
  - Which processes produce the smallest variances?
  - What are the customer attrition rates for various quality levels?

Directed Knowledge Discovery-2

- Identify sources of preclassified data
- Prepare data
  - Data cleansing
  - Compute additional fields
- Build and train a computer model
- Evaluate the computer model
- Learn from outliers

Undirected Knowledge Discovery-1

- Generate ideas to be evaluated using directed KD and hypothesis testing
- E.g., “Which defects occur together?” rather than “Which defects occur with cracks?”

Undirected Knowledge Discovery-2

- Identify the sources of data
- Prepare the data
- Build and train computer model
- Evaluate the computer model
- Apply computer model to new data
- Identify targets for directed KD
- Generate new hypotheses to test

Information Systems Quality

Why IS Quality?-1

- IS is vital to serving customer
- IS systems are integrated into everyone's work
- IS systems are very costly to purchase
- IS systems are very costly to operate
- IS systems are becoming more visible to customers

Why IS Quality?-2

- IS Quality stinks
  - Systems are unreliable
  - IS systems are user-hostile
  - Data are hard to find
  - Data are incorrect
  - Data are missing
  - Data are difficult to interpret

Real Life Examples

Problem	Cost
In the past year there were several instances where web site customers downloaded fully functional units rather than demo versions	$3.1 million in lost revenues
12% of email responses to customer inquiries are undeliverable	$ Unknown. 12,000 undeliverables/month
24% of all customer web site issues are not resolved	$ Unknown. 800 unresolved issues/month

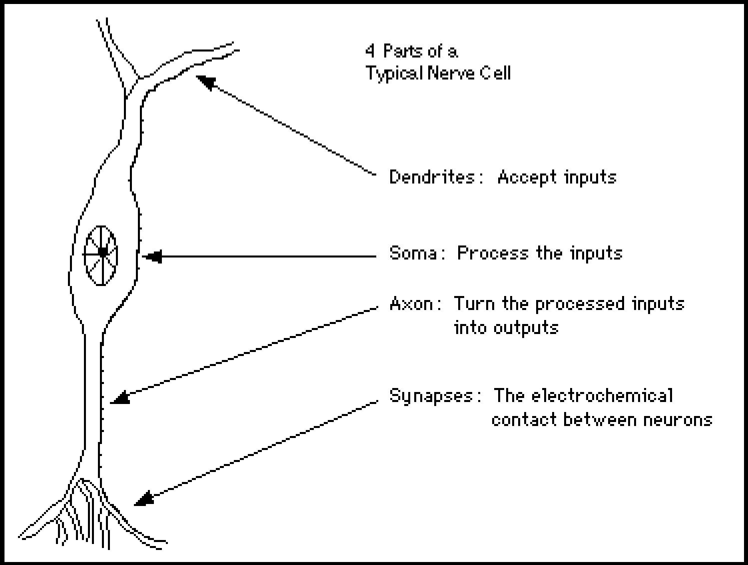
How Can Six Sigma Help?

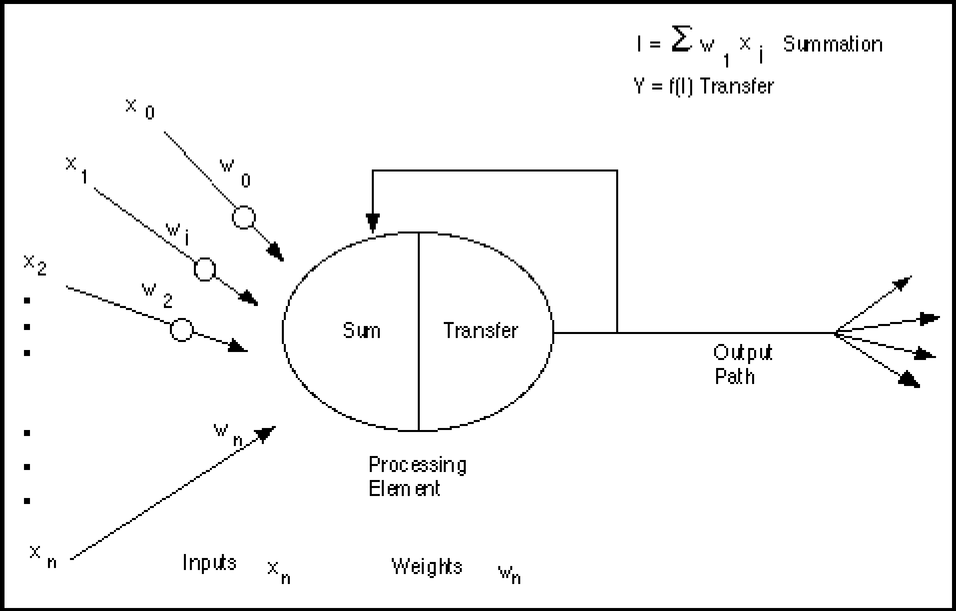
- Apply tried-and-true quality technology to IS
  - Determine cost of poor quality
  - Establish basic quality systems
    - Apply to procurement, design,, and development
  - Apply Six Sigma methods to improving the IS process (DMAIC)

Virtual DOE using Data Mining and Artificial Neural Networks

What Are Artificial Neural Networks?-1

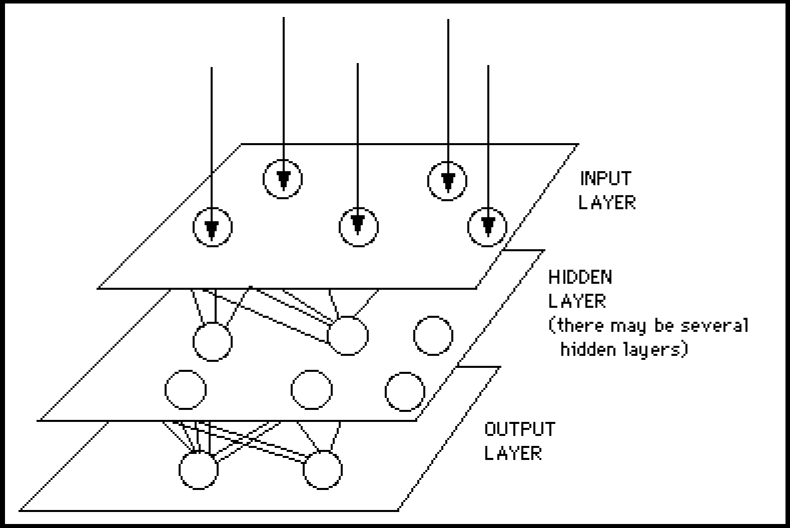
- Computer models based on *very* crude approximations to the neural connections of brains



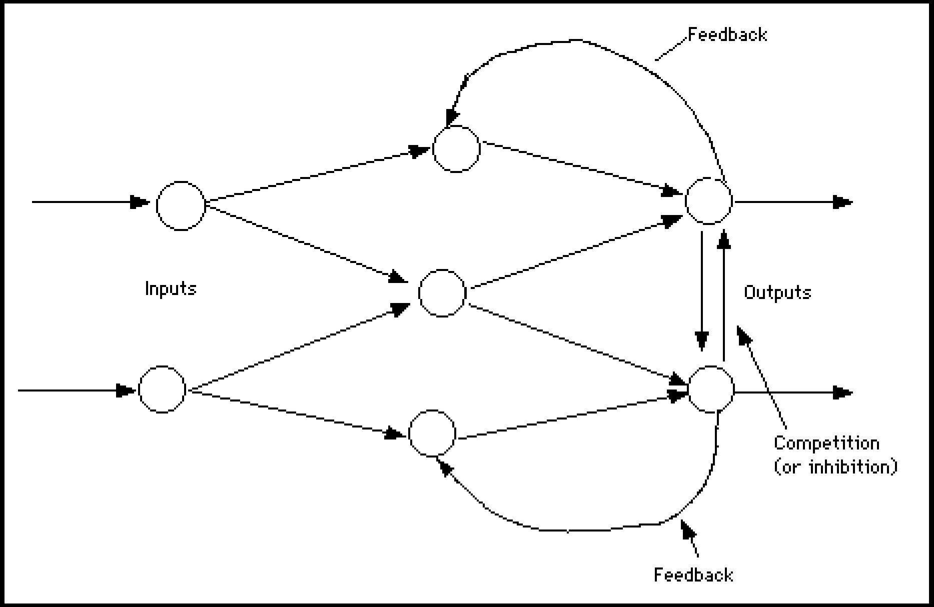


What Are Artificial Neural Networks?-2

Neural Network



Neural Network
With Feedback



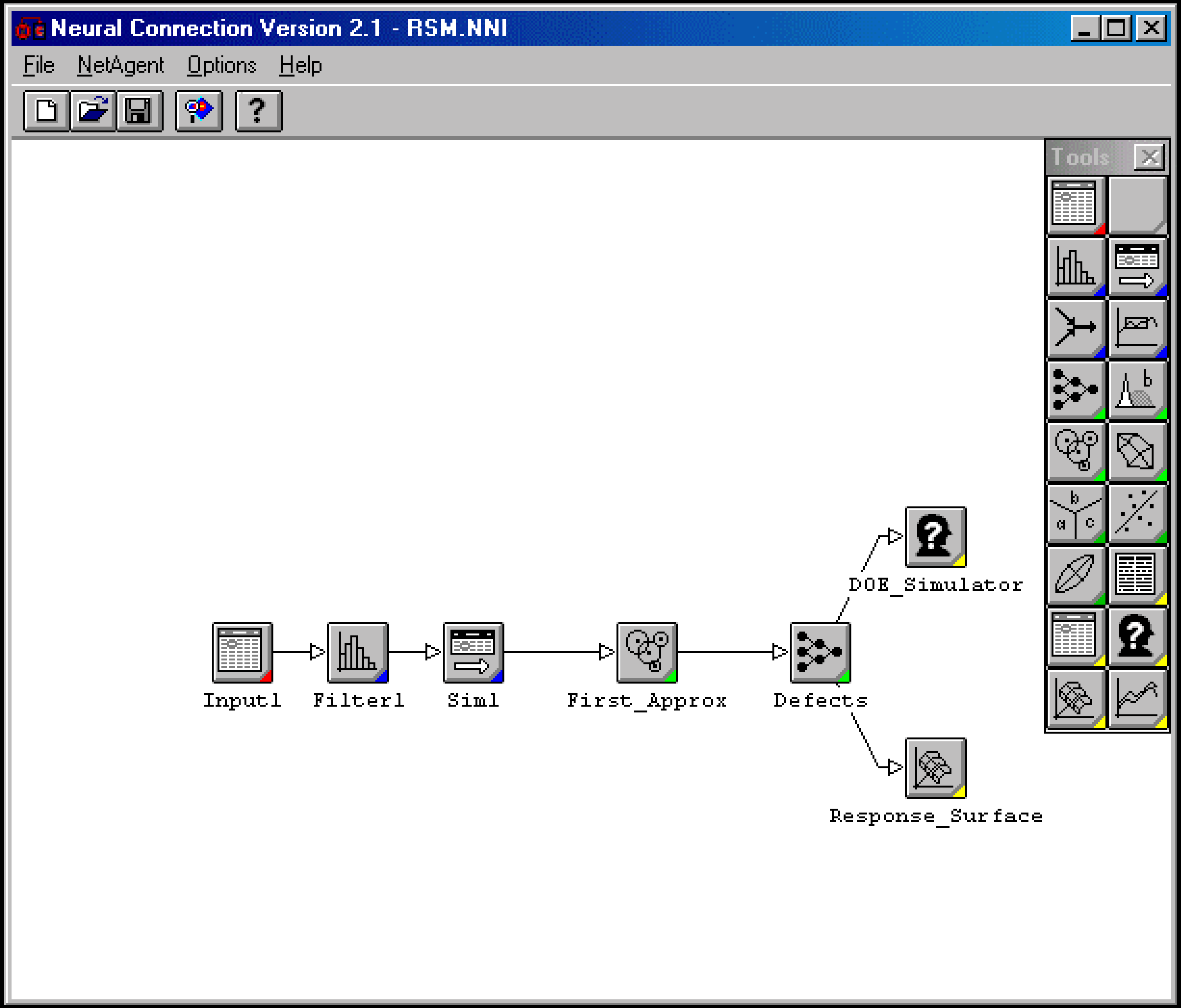
What Is Needed to Use Neural Nets?

- Characteristics of candidate problems
  - The inputs are well understood
    - You know what's important, but not how to model it
  - The output is well understood
    - You know what you are trying to predict
  - Experience is available
    - Examples will be used to train the network

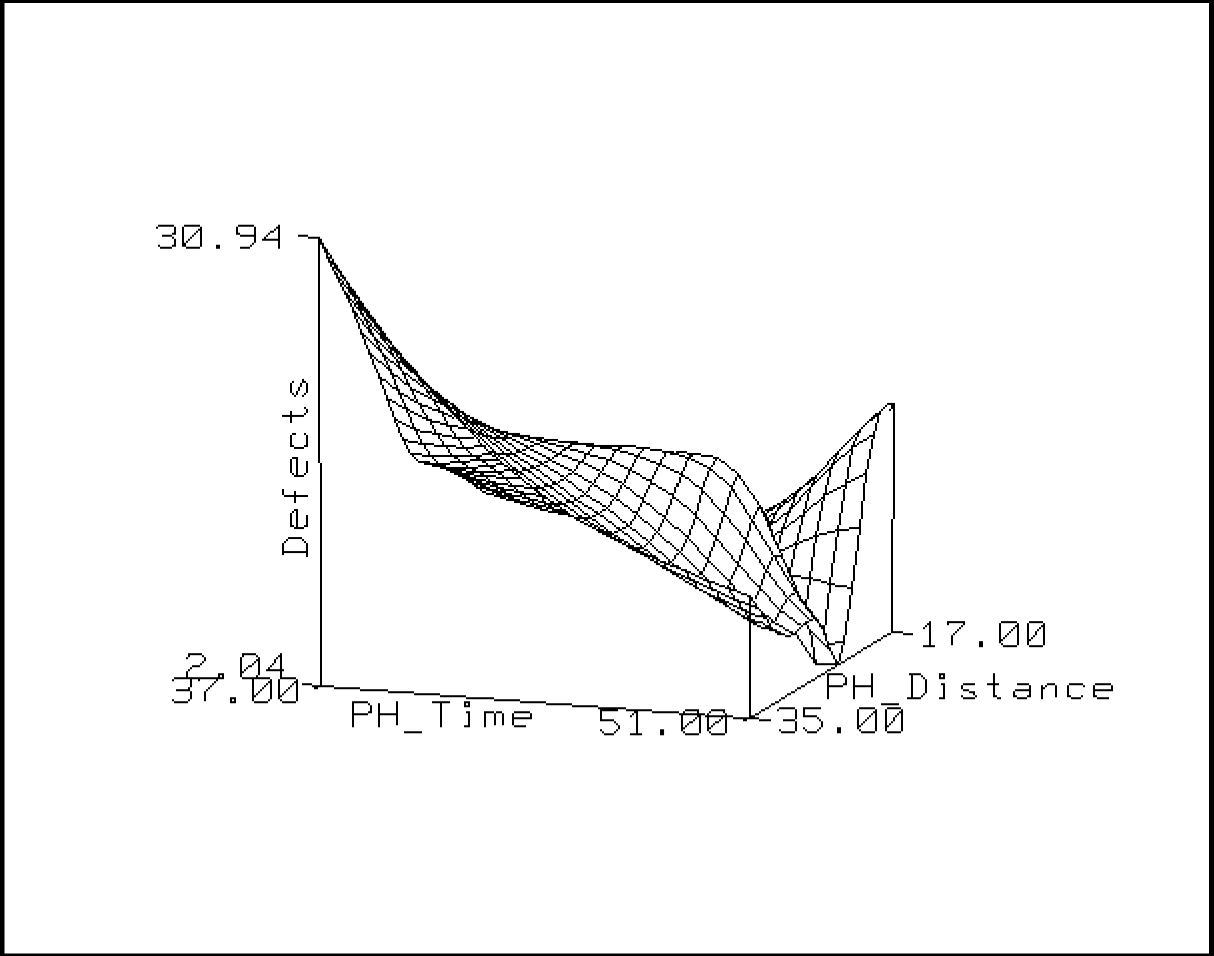
The Data

PH_Time	PH	Distan	Defects
38		22.5	15
40		20.0	13
40		25.0	16
45		17.5	15
45		22.5	5
45		26.0	11
50		20.0	12
42		22.5	10
50		25.0	3
42		22.0	11
46		22.0	4
55		25.0	4
55		21.0	17
55		25.0	15
50		24.0	3
49		25.0	3
57		37.0	10
35		25.0	20
45		37.5	17
30		20.0	27
30		22.5	33
30		25.0	37
30		27.5	50
30		37.5	57
50		20.0	13
50		22.5	5
50		25.0	3
50		30.0	5
50		14.0	12
50		37.5	14
50		45.0	16
50		50.0	40
60		20.0	35
60		25.0	18
60		37.5	12

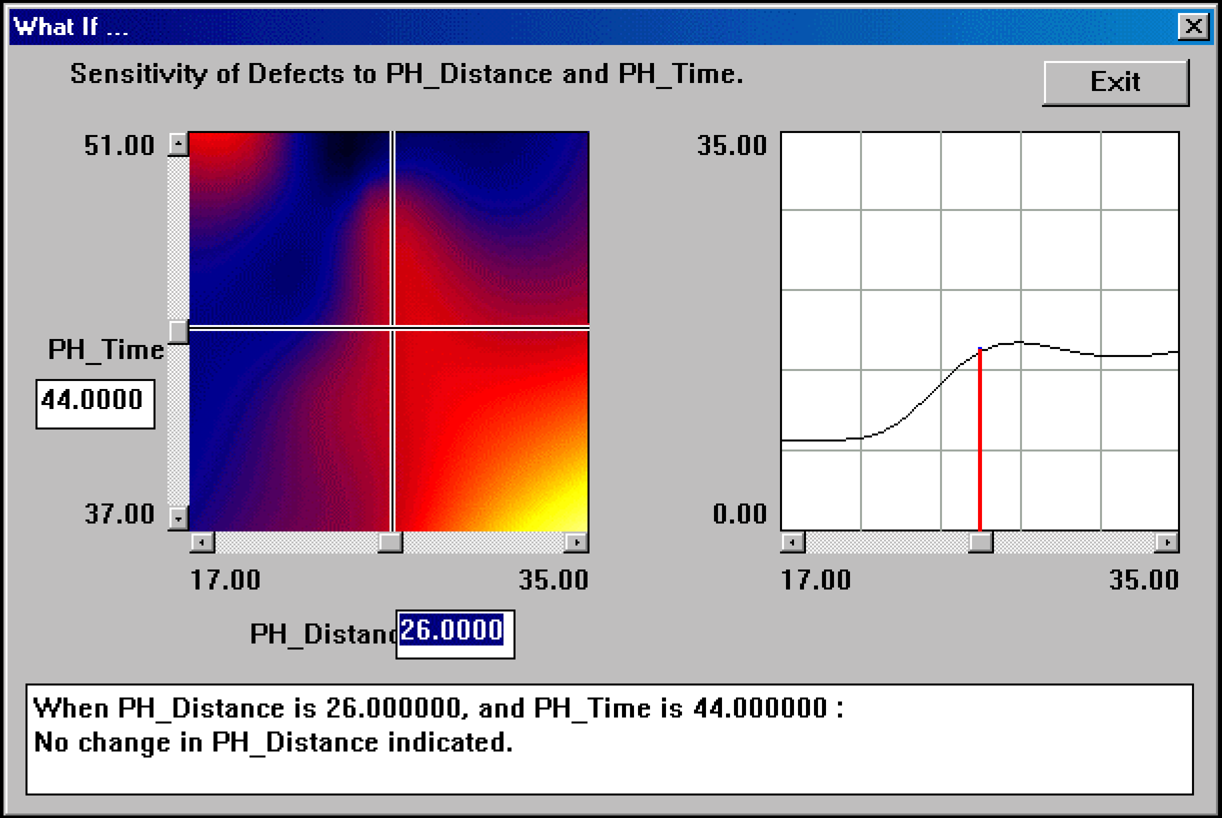
The Model


Neural Connection Version 2.1 - RSM.NNI
File
NetAgent
Options
Help
Tools
Input1
Filter1
Sim1
First_Approx
Defects
DOE_Simulator
Response_Surface

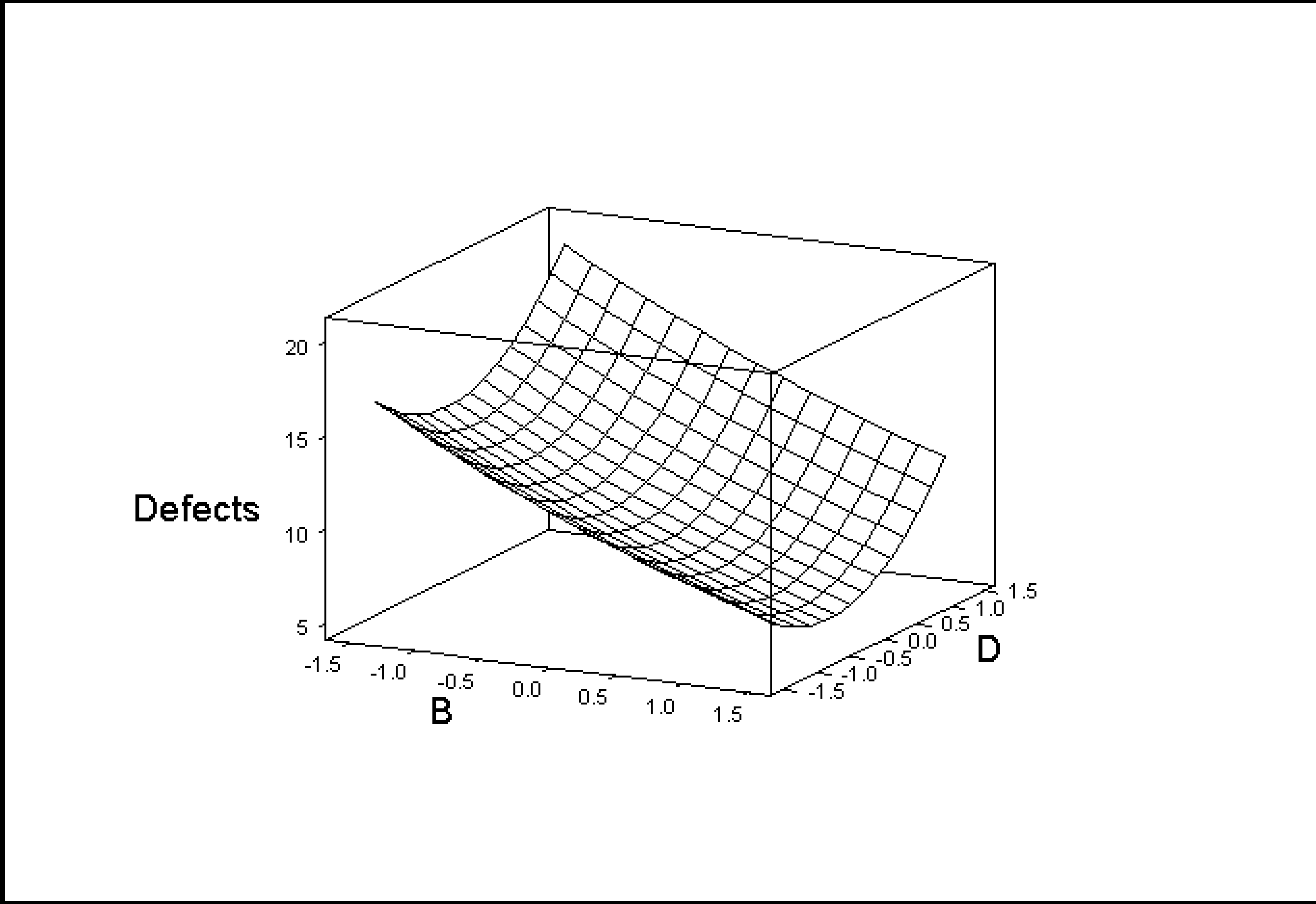
The Response Surface Modeled by the Neural Net



What If? Analysis & VDOE Can Get You in the Neighborhood

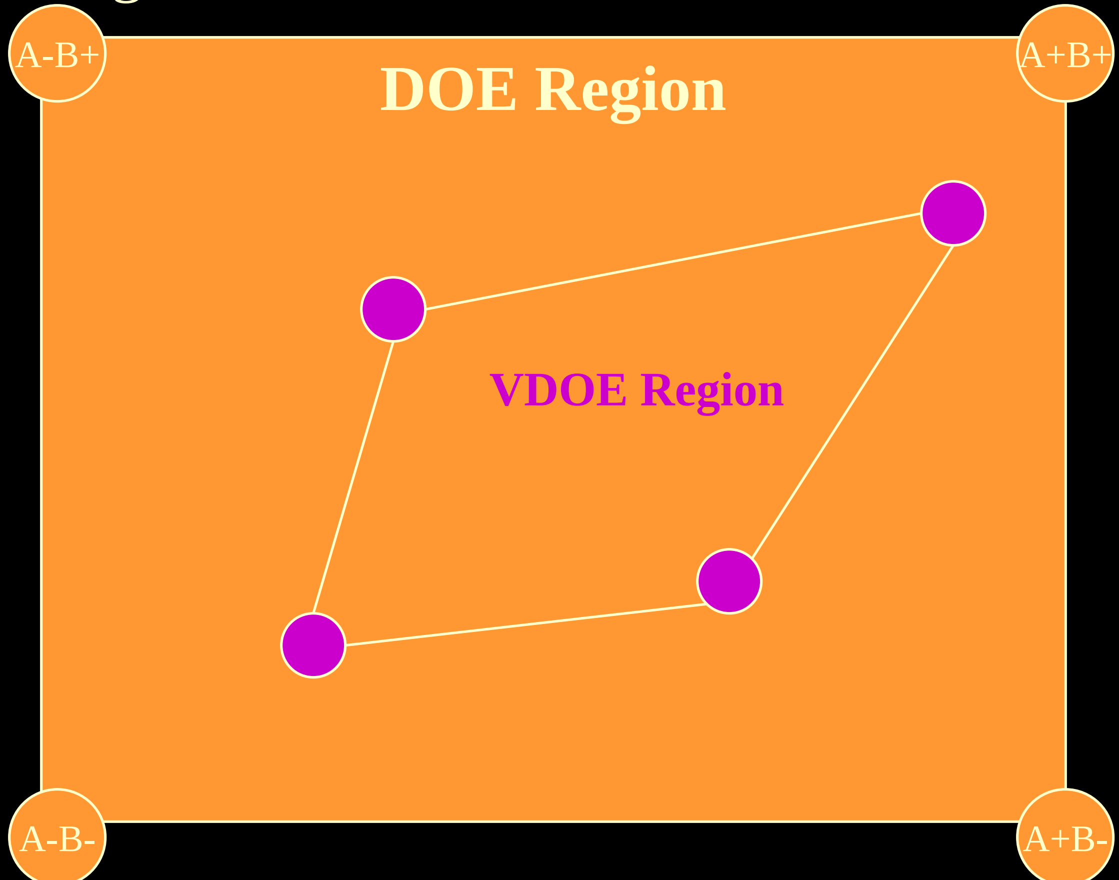


The Response Surface Found Using DOE

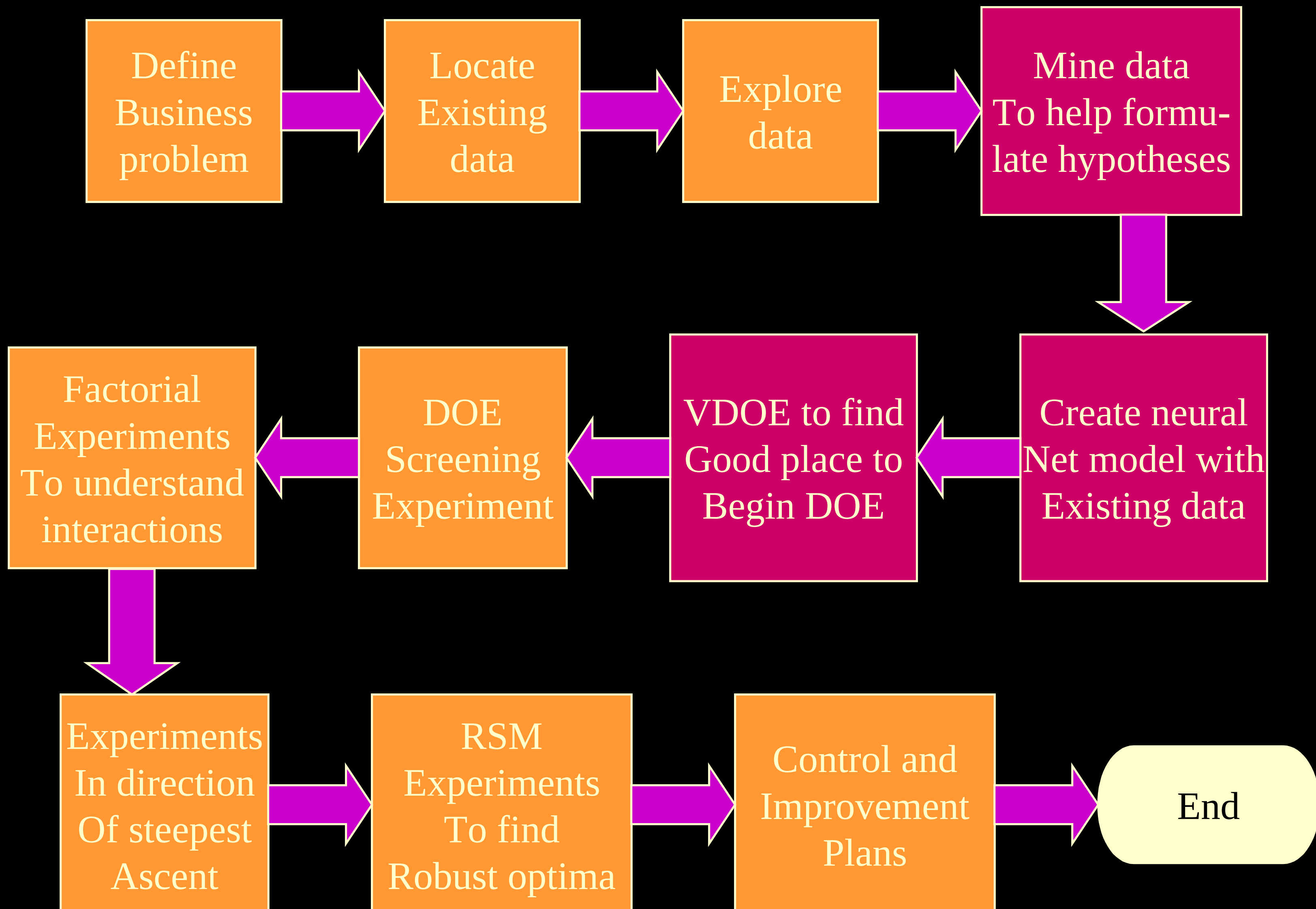


Data Mining Cautions

- Data quality issues (errors, relevance)
- Data availability issues
  - Missing Xs, non-orthogonal, unbalanced
- Data range issues



Approach


New
Traditional
Define Business problem
Locate Existing data
Explore data
Mine data To help formu-late hypotheses
Create neural Net model with Existing data
VDOE to find Good place to Begin DOE
DOE Screening Experiment
Factorial Experiments To understand interactions
Experiments In direction Of steepest Ascent
RSM Experiments To find Robust optima
Control and Improvement Plans
End

Example

- Define

  Too many parts don't meet the customer minimum requirement for test to failure for an injection molded plastic part.

- Measure

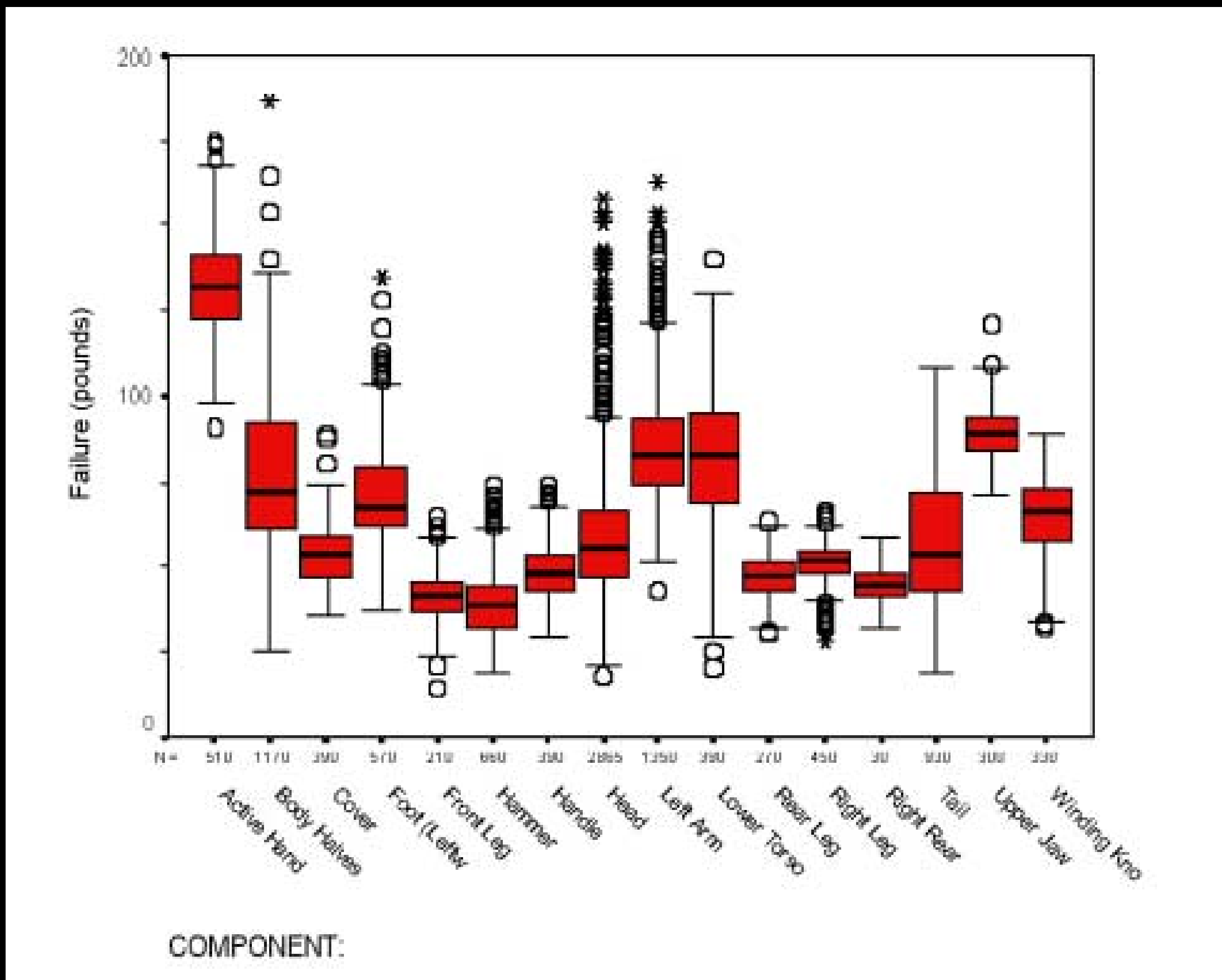
  There are two CTQs: average test to failure force, and standard deviation.

Analysis

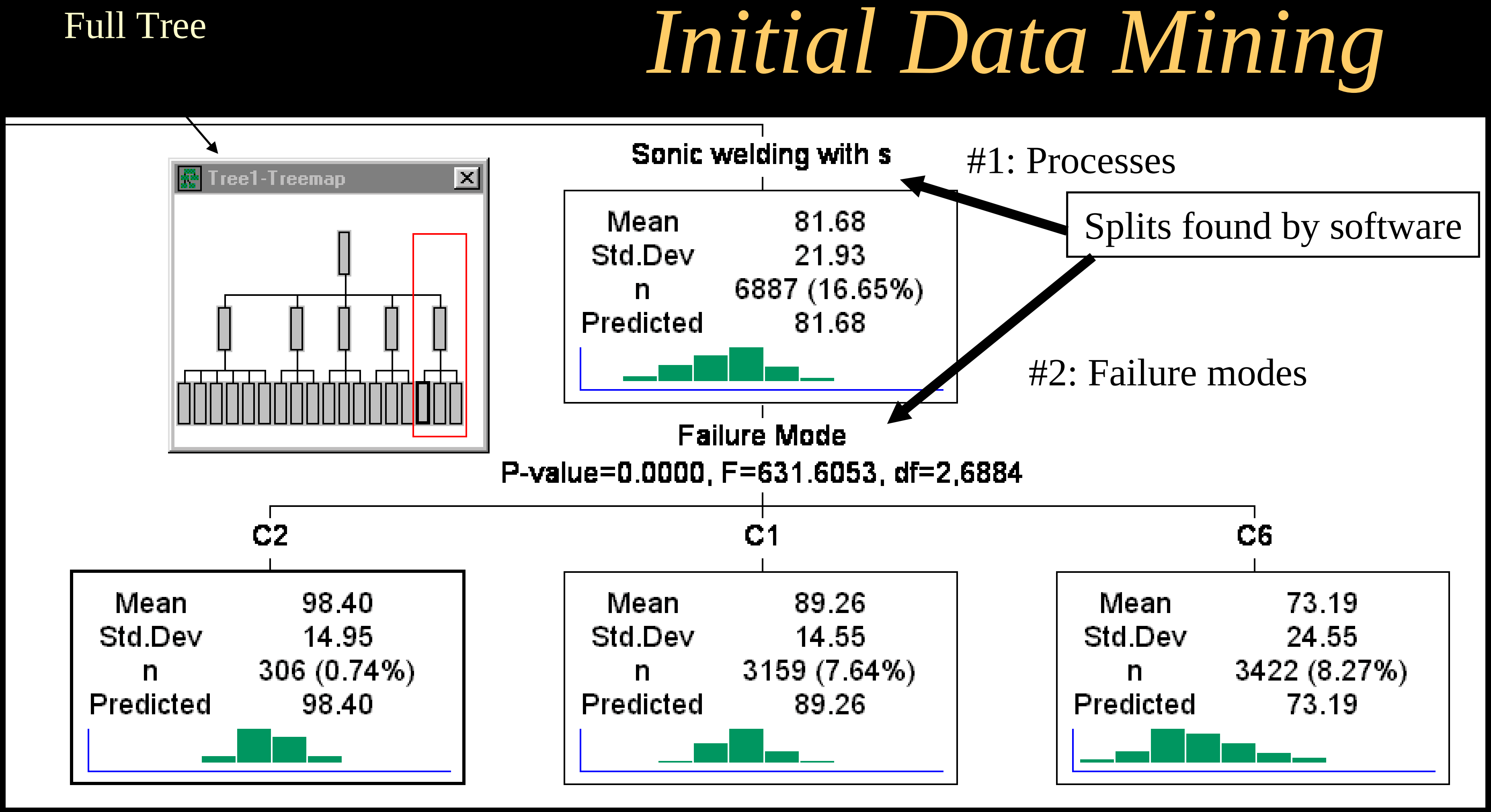
- Analyze

  Perform data mining to identify promising hypotheses

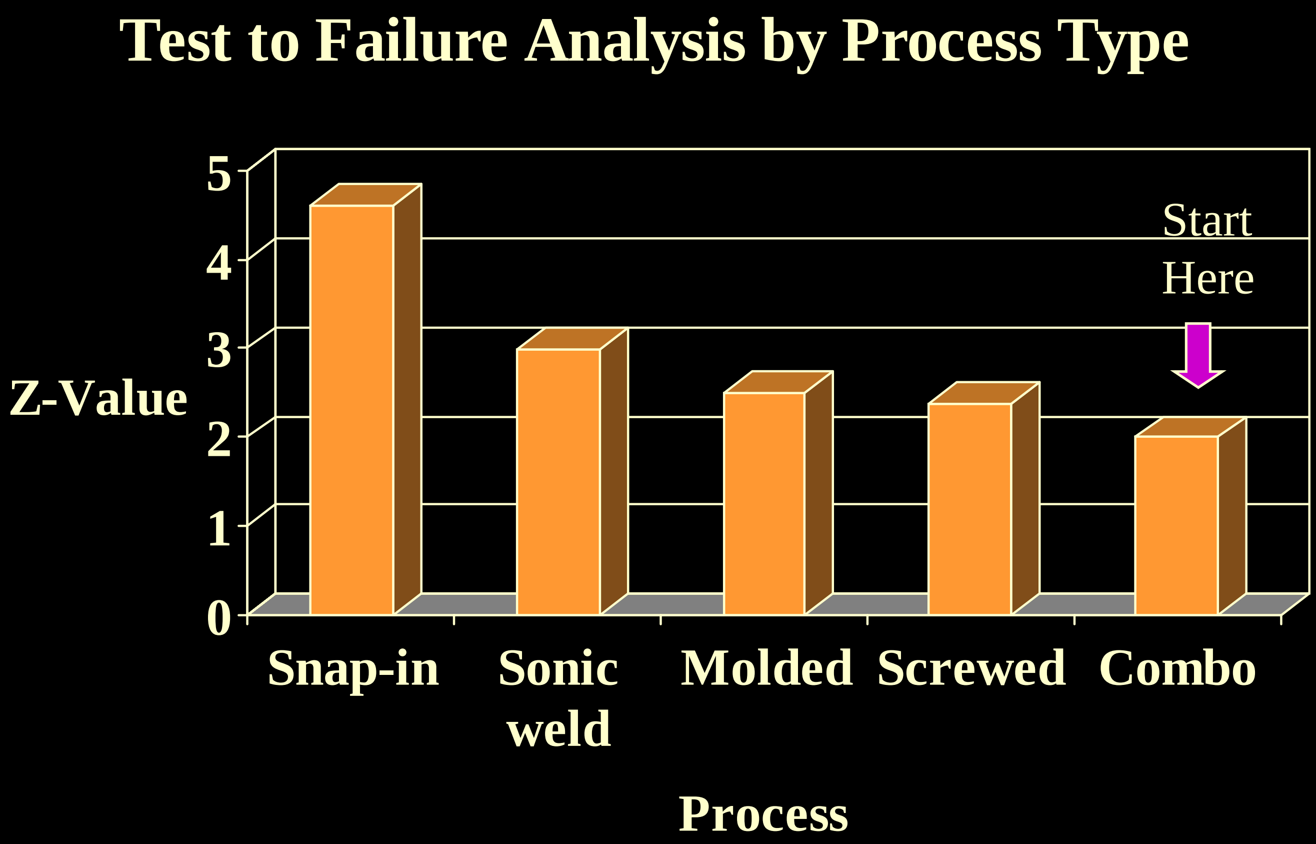
Drill Down and Explore

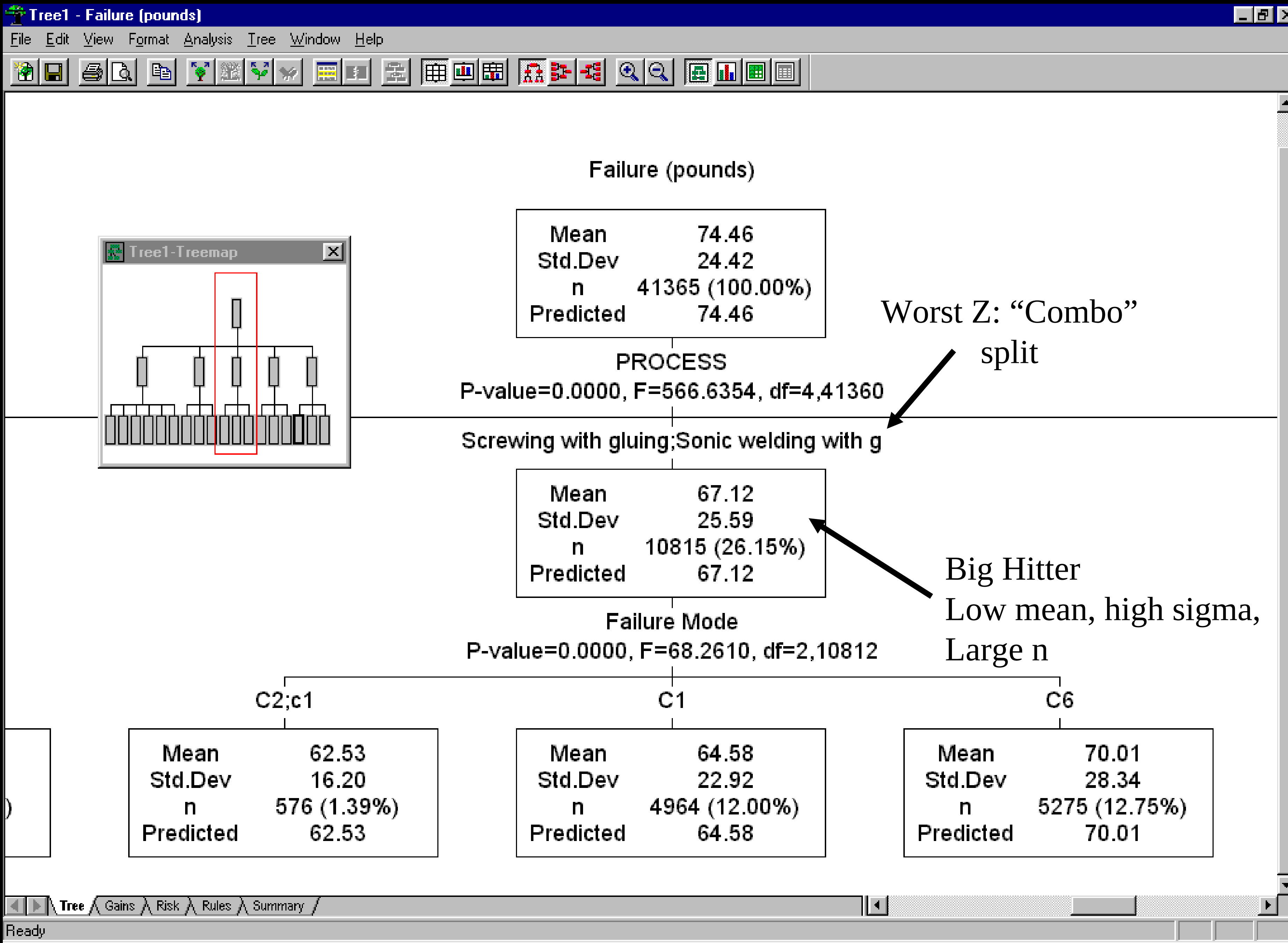


Initial Data Mining



Prioritize


Test to Failure Analysis by Process Type
Z-Value
5
4
3
2
1
0
Start
Here
Snap-in
Sonic weld
Molded
Screwed
Combo
Process


Tree1 - Failure (pounds)
File Edit View Format Analysis Tree Window Help
Tree1-Treemap
Failure (pounds)
Mean 74.46
Std.Dev 24.42
n 41365 (100.00%)
Predicted 74.46
PROCESS
P-value=0.0000, F=566.6354, df=4,41360
Screwing with gluing;Sonic welding with g
Mean 67.12
Std.Dev 25.59
n 10815 (26.15%)
Predicted 67.12
Failure Mode
P-value=0.0000, F=68.2610, df=2,10812
C2;c1
Mean 62.53
Std.Dev 16.20
n 576 (1.39%)
Predicted 62.53
C1
Mean 64.58
Std.Dev 22.92
n 4964 (12.00%)
Predicted 64.58
C6
Mean 70.01
Std.Dev 28.34
n 5275 (12.75%)
Predicted 70.01
Worst Z: “Combo” split
Big Hitter
Low mean, high sigma,
Large n
Tree Gains Risk Rules Summary
Ready

Extract data from the node of interest

Tree1 - Data : Node 31

File Edit View Help

	strength		design	componen	failmode	process
1	80		Right Leg	Foot (Leftw	C1	Screwing with gluing
2	62		Right Leg	Foot (Leftw	C1	Screwing with gluing
3	61		Right Leg	Foot (Leftw	C1	Screwing with gluing
4	63		Right Leg	Foot (Leftw	C1	Screwing with gluing
5	61		Right Leg	Foot (Leftw	C1	Screwing with gluing
6	90		Right Leg	Foot (Leftw	C1	Screwing with gluing
7	65		Right Leg	Foot (Leftw	C1	Screwing with gluing
8	69		Right Leg	Foot (Leftw	C1	Screwing with gluing
9	80	Confidential	Right Leg	Foot (Leftw	C1	Screwing with gluing
10	59		Right Leg	Foot (Leftw	C1	Screwing with gluing
11	74		Right Leg	Foot (Leftw	C1	Screwing with gluing
12	78		Right Leg	Foot (Leftw	C1	Screwing with gluing
13	78		Right Leg	Foot (Leftw	C1	Screwing with gluing
14	113		Right Leg	Foot (Leftw	C1	Screwing with gluing
15	77		Right Leg	Foot (Leftw	C1	Screwing with gluing
16	61		Right Leg	Foot (Leftw	C1	Screwing with gluing
17	70		Right Leg	Foot (Leftw	C1	Screwing with gluing

Note: Program also creates SQL query to extract from large databases

Hypothesize and experiment

- $Y = f(X_1, X_2, X_3)$
- Strength = *f*(Cross Sectional Area, Material, radius)
- Train neural net and conduct VDOE to find starting point
- DOEs
- Prepare control and improvement plans

Questions?