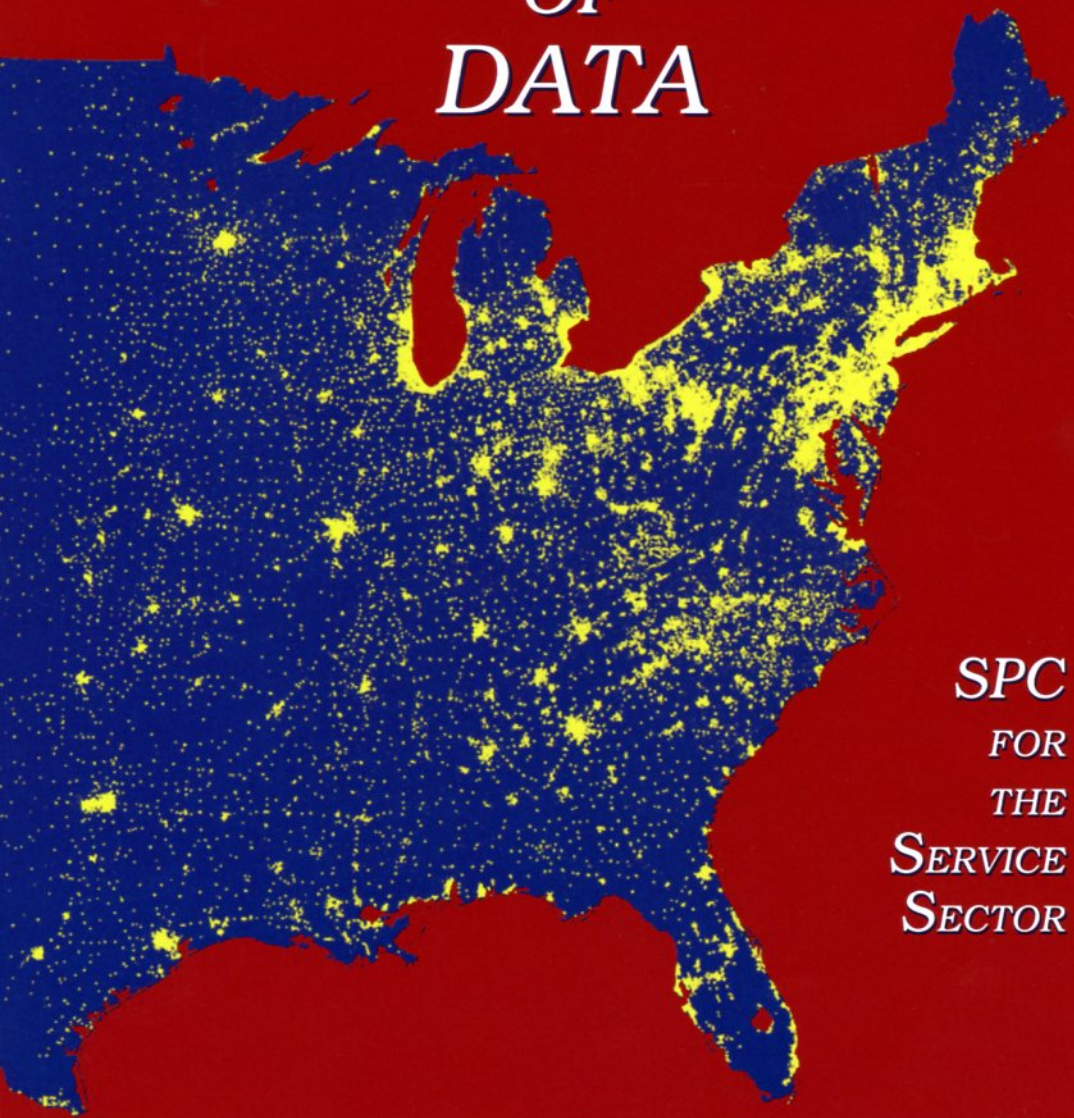
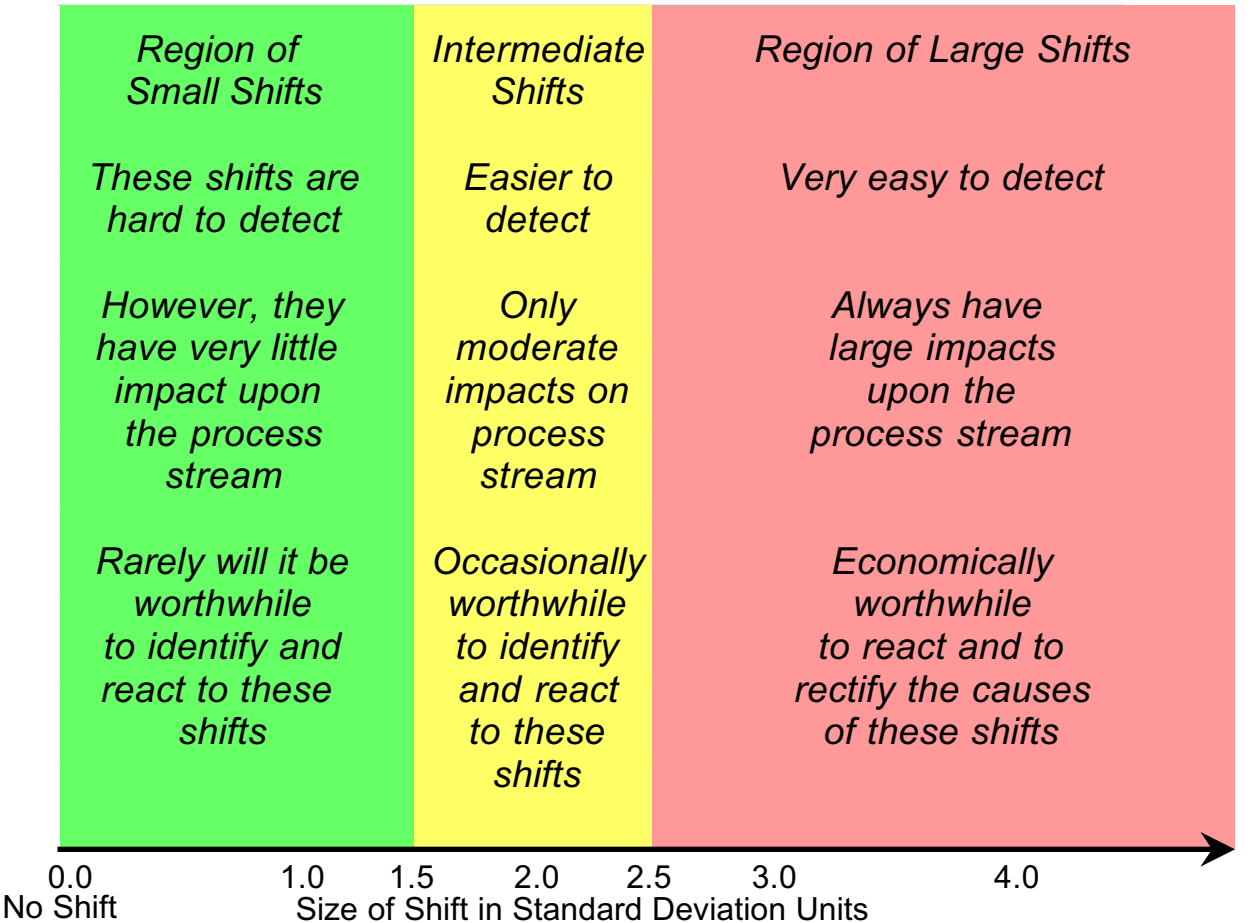
MAKING SENSE OF DATA

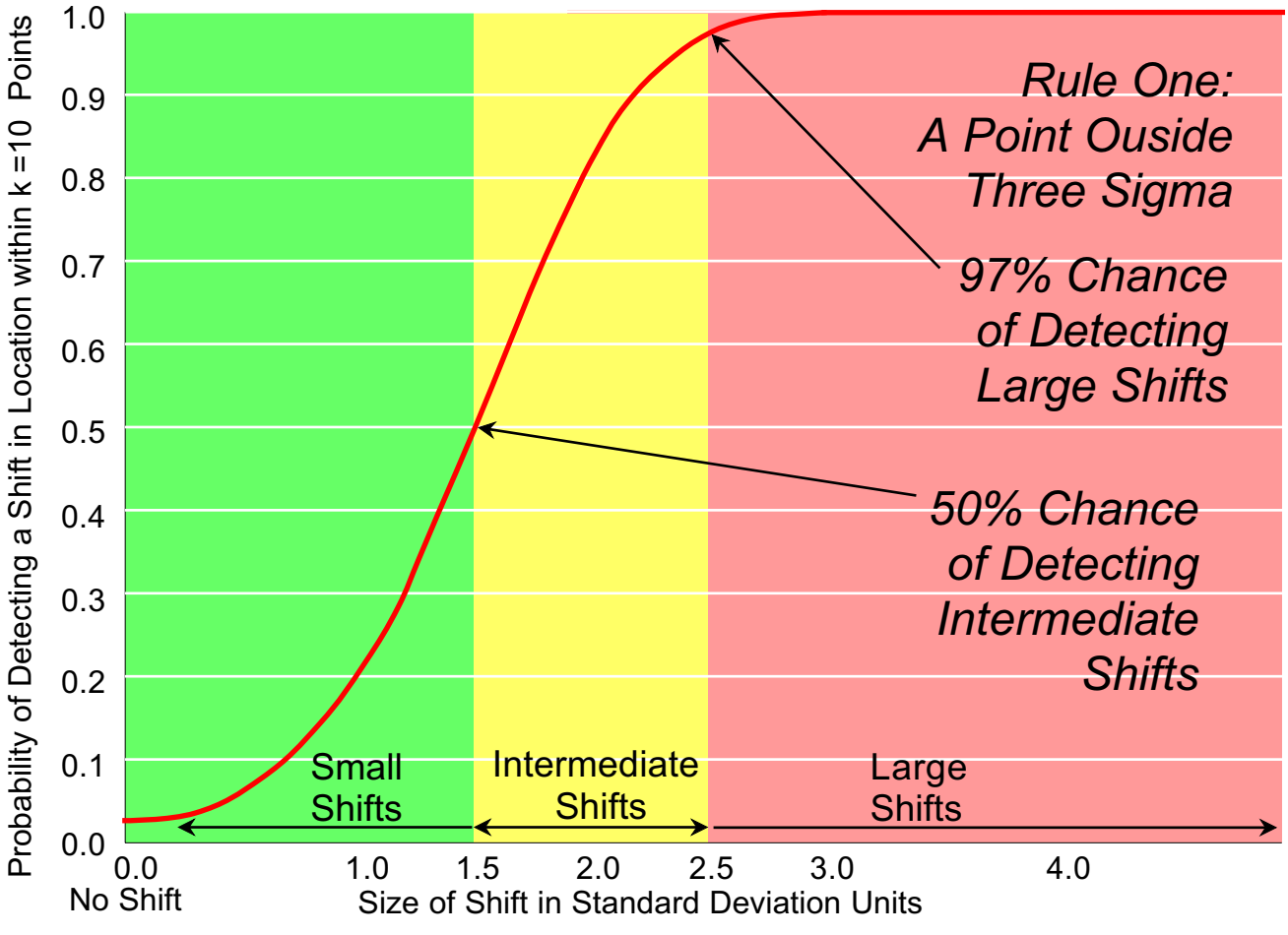


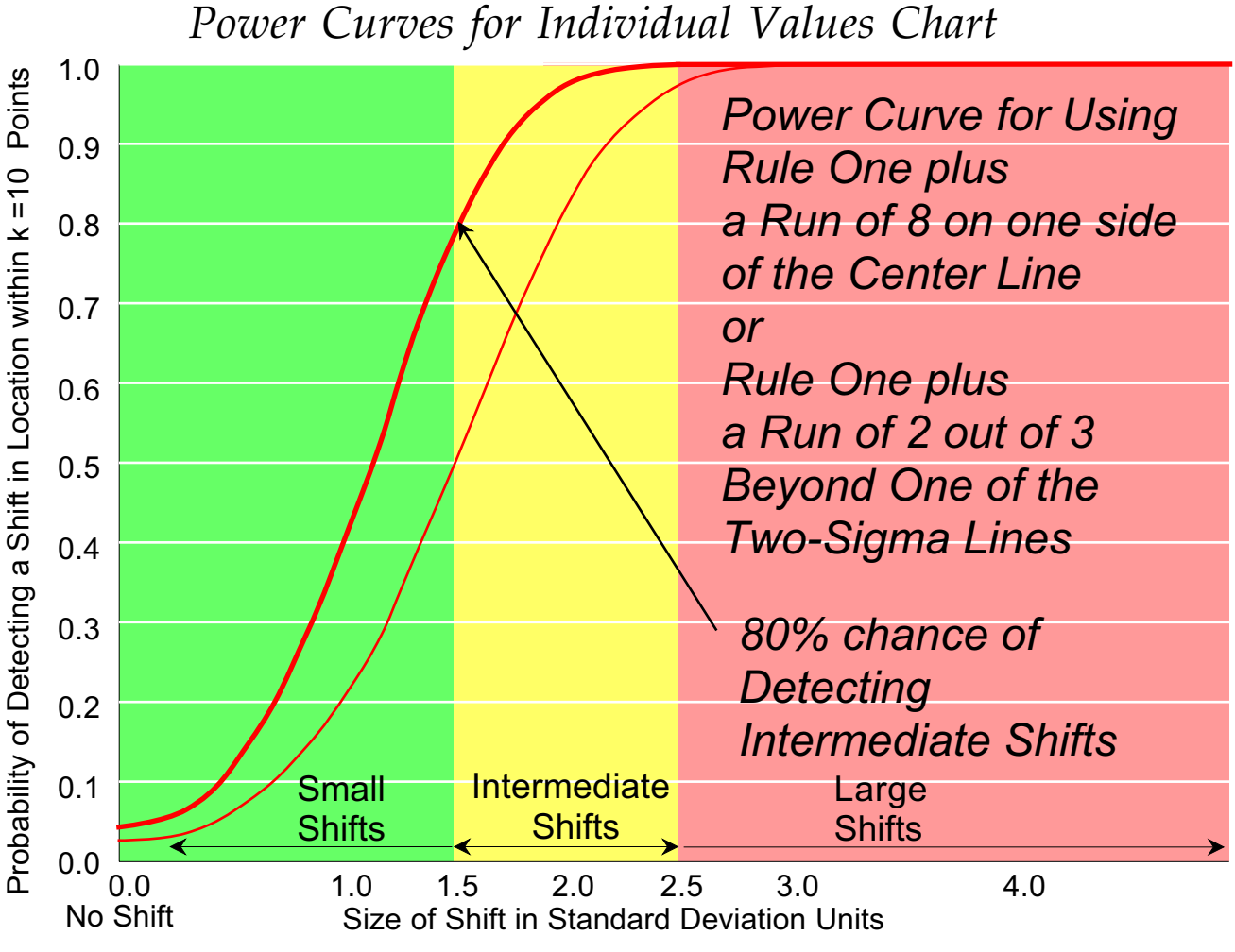
SPC FOR THE SERVICE SECTOR

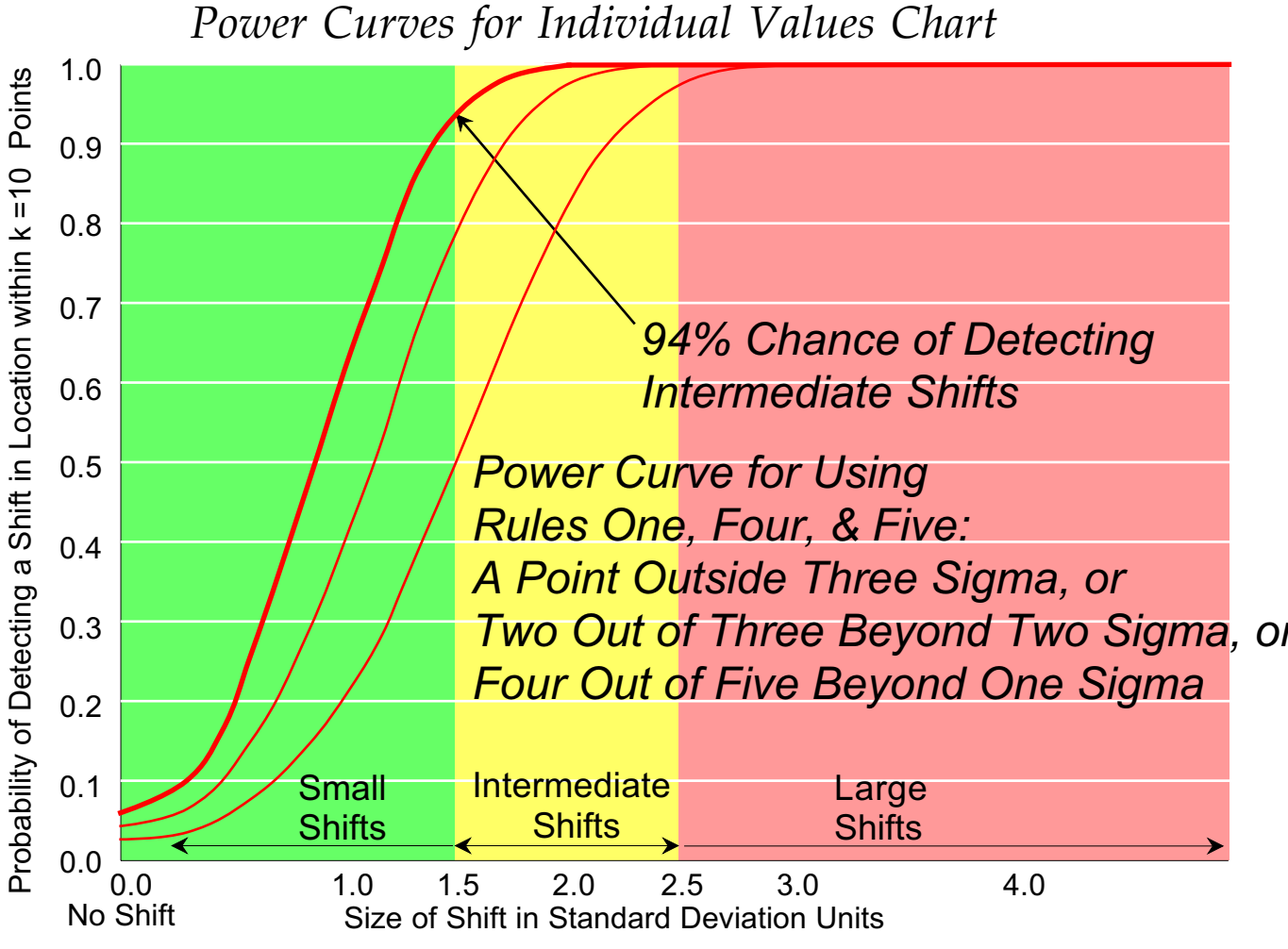
DONALD J. WHEELER

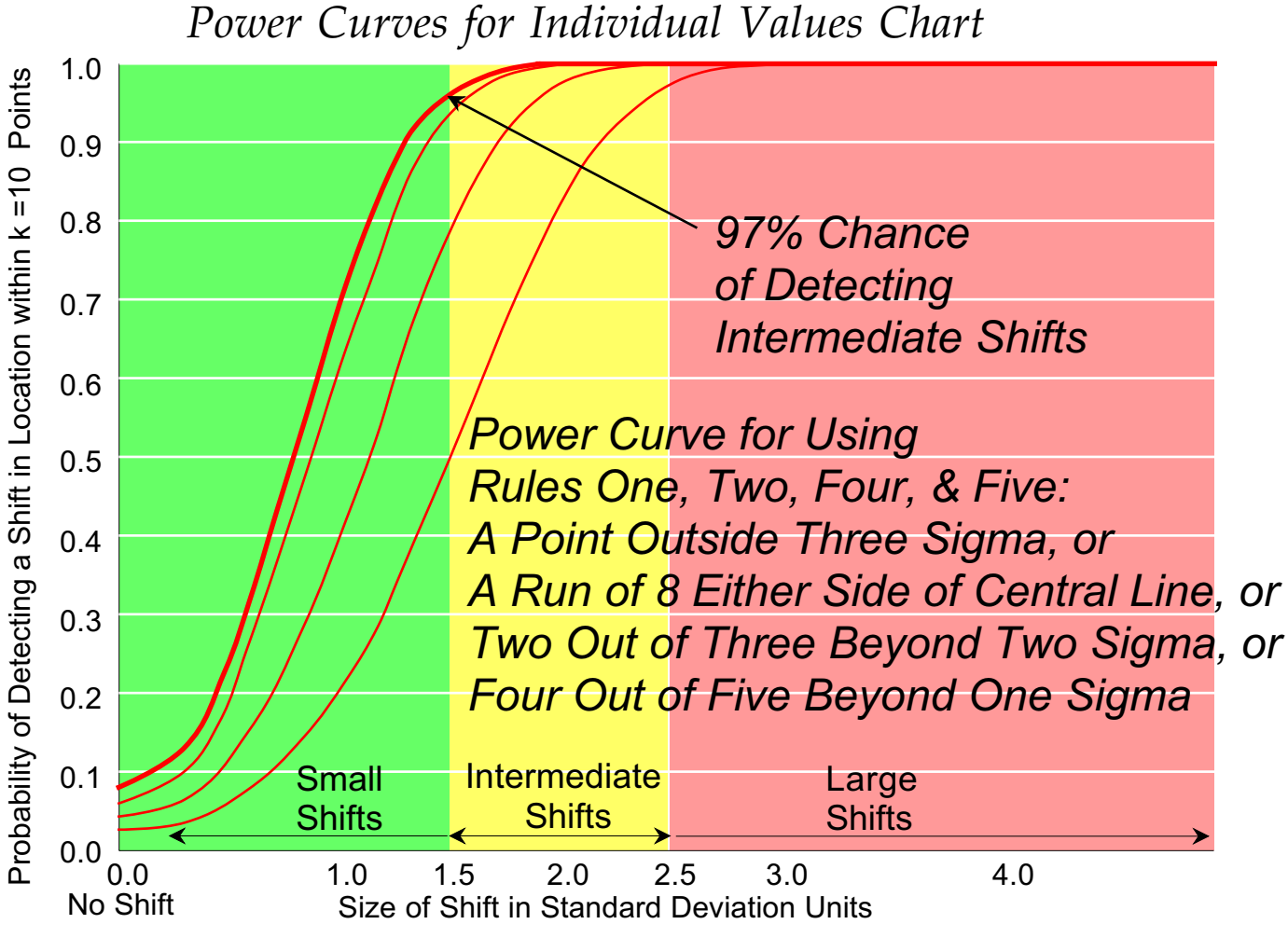
Different Size Shifts Have Different Impacts Upon Process




Power Curves for Individual Values Chart
Probability of Detecting a Shift in Location within k =10 Points
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Rule One:
A Point Ouside
Three Sigma
97% Chance
of Detecting
Large Shifts
50% Chance
of Detecting
Intermediate
Shifts
Small
Shifts
Intermediate
Shifts
Large
Shifts
0.0
No Shift
1.0
1.5
2.0
2.5
3.0
4.0
Size of Shift in Standard Deviation Units


Power Curves for Individual Values Chart
Probability of Detecting a Shift in Location within k =10 Points
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Power Curve for Using Rule One plus a Run of 8 on one side of the Center Line or Rule One plus a Run of 2 out of 3 Beyond One of the Two-Sigma Lines
80% chance of Detecting Intermediate Shifts
Small Shifts
Intermediate Shifts
Large Shifts
0.0
No Shift
1.0
1.5
2.0
2.5
3.0
4.0
Size of Shift in Standard Deviation Units


Power Curves for Individual Values Chart
Probability of Detecting a Shift in Location within k =10 Points
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
94% Chance of Detecting Intermediate Shifts
Power Curve for Using Rules One, Four, & Five: A Point Outside Three Sigma, or Two Out of Three Beyond Two Sigma, or Four Out of Five Beyond One Sigma
Small Shifts
Intermediate Shifts
Large Shifts
0.0
No Shift
1.0
1.5
2.0
2.5
3.0
4.0
Size of Shift in Standard Deviation Units


Power Curves for Individual Values Chart
Probability of Detecting a Shift in Location within k =10 Points
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
97% Chance
of Detecting
Intermediate Shifts
Power Curve for Using
Rules One, Two, Four, & Five:
A Point Outside Three Sigma, or
A Run of 8 Either Side of Central Line, or
Two Out of Three Beyond Two Sigma, or
Four Out of Five Beyond One Sigma
Small
Shifts
Intermediate
Shifts
Large
Shifts
0.0
No Shift
1.0
1.5
2.0
2.5
3.0
4.0
Size of Shift in Standard Deviation Units

Making Sense of Data Rules Used	Detect Large Shifts >2.5 Sigma	Detect Intermediate Shifts > 1.5 Sigma
One	97%	50%
One & Two	100%	80%
One & Four	100%	80%
One, Four, & Five	100%	94%
One, Two, Four, & Five	100%	97%

When looking for signals of change that are large enough
to guarantee that it will be economical
to discover the assignable causes of exceptional variation
and to make these assgnable causes
part of the set of controlled process inputs,
Rule One is sufficient.

www.spcpress.com/pdf/DJW350.pdf

www.spcpress.com/pdf/DJW321.pdf

www.spcpress.com/pdf/DJW322.pdf

Why Susie Can Breathe

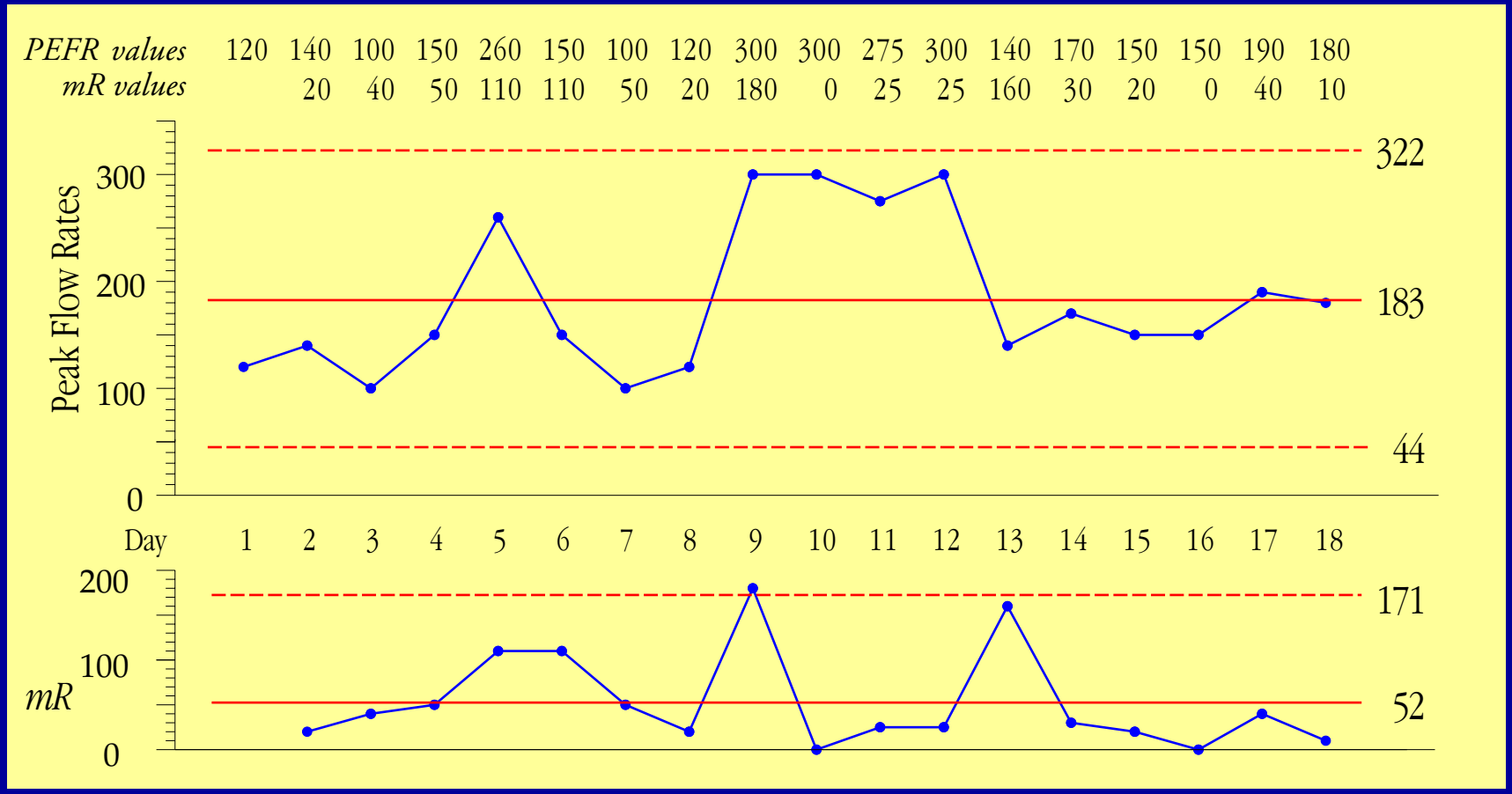
An 11-year old with severe asthma
who had been hospitalized three times in the past nine months
was brought to Dr. Peter Boggs, an allergist.

After the usual tests for allergens,
Dr. Boggs gave this girl the appropriate prescriptions
and a PEFR gauge to use.

Upon the next office visit Dr. Boggs plotted
the A.M. pre-medication PEFR scores
on an *XmR* chart.

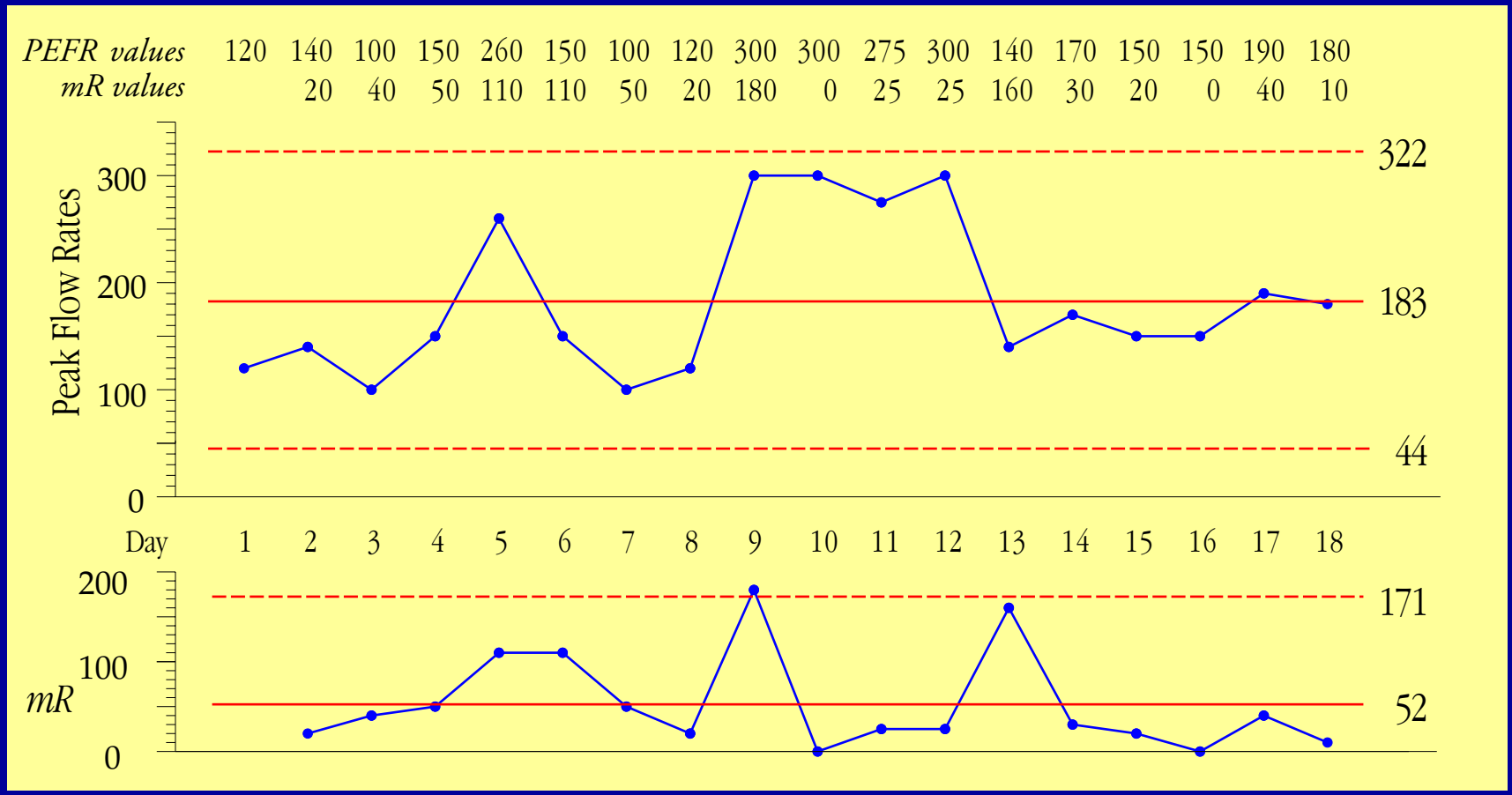
Why Susie Can Breathe

A.M. pre-medication PEFR scores



Why Susie Can Breathe

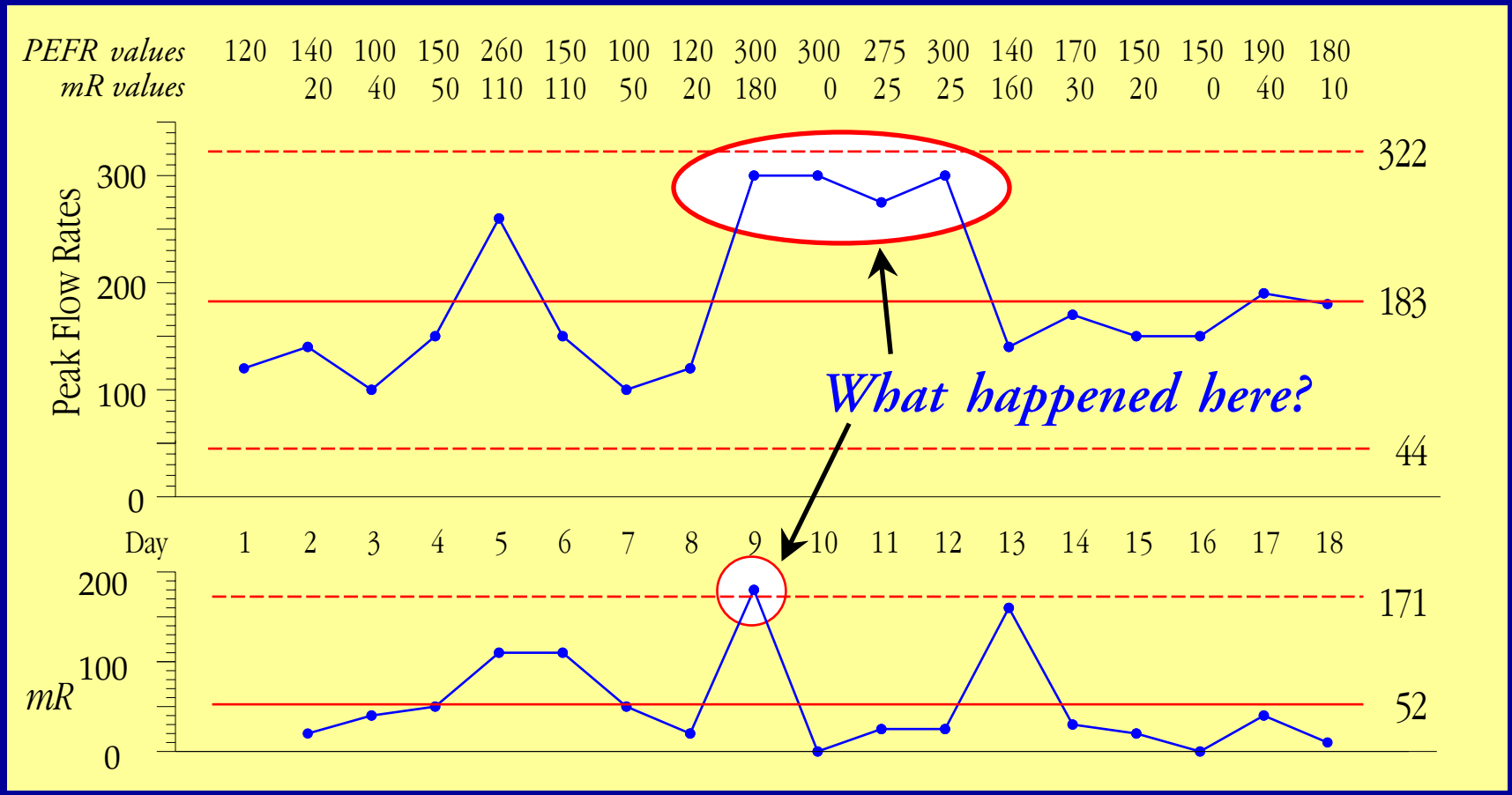
Upper limit of 322 is consistent with her predicted best of 310.



Lower limit of 44 is consistent with medical history.
And average of 183 is only 60 percent of predicted best.

Why Susie Can Breathe

But the interesting part of this chart is the run near the limit on Days 9 to 12.

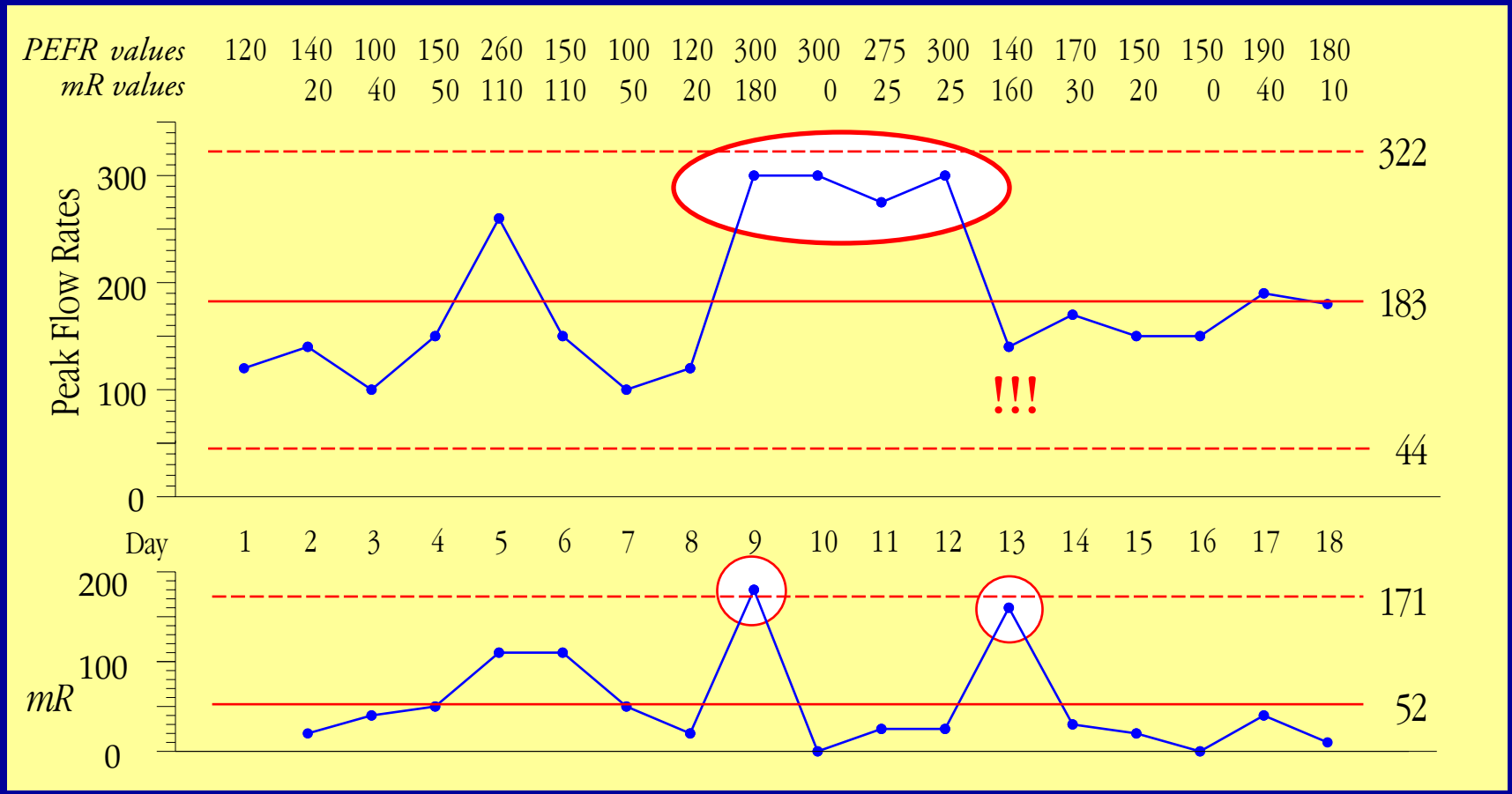


"That's when she went to her aunt's for the weekend."

Why Susie Can Breathe

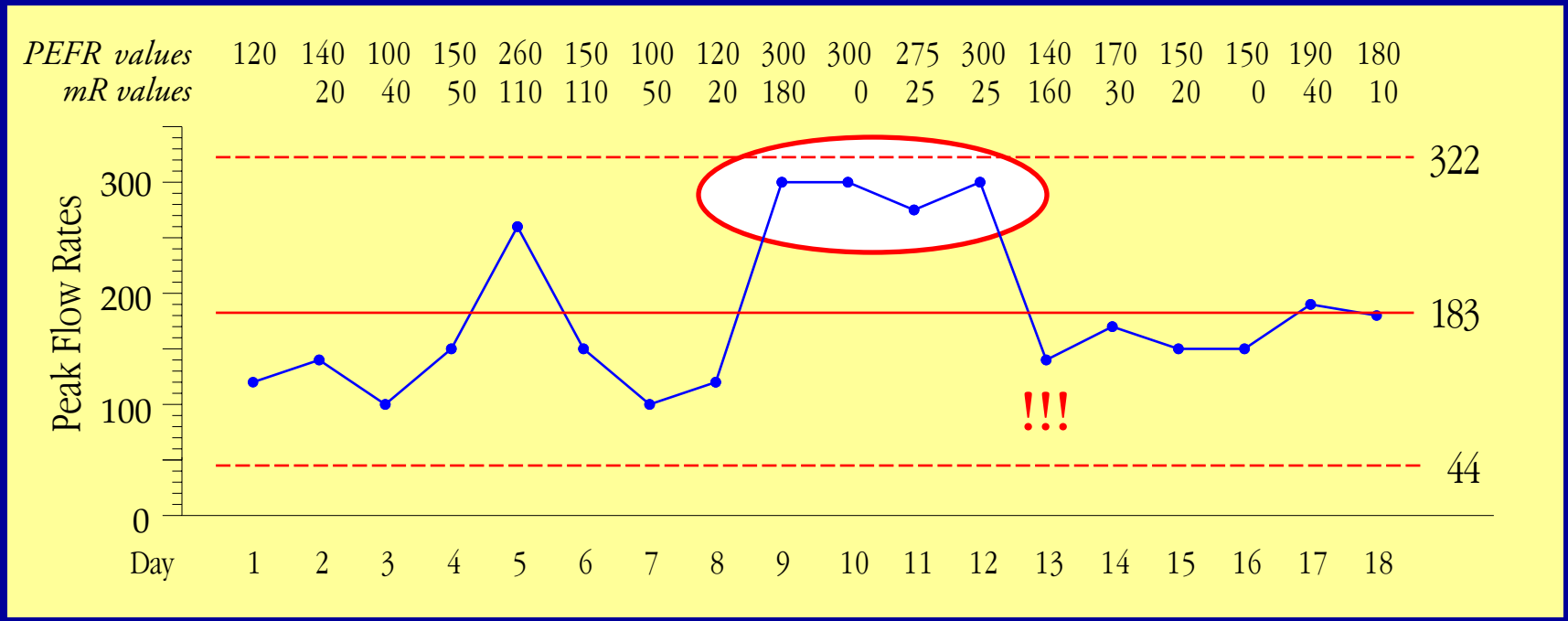
"And then she returned home on day 13 and got sick again?"

"Yes. She did."



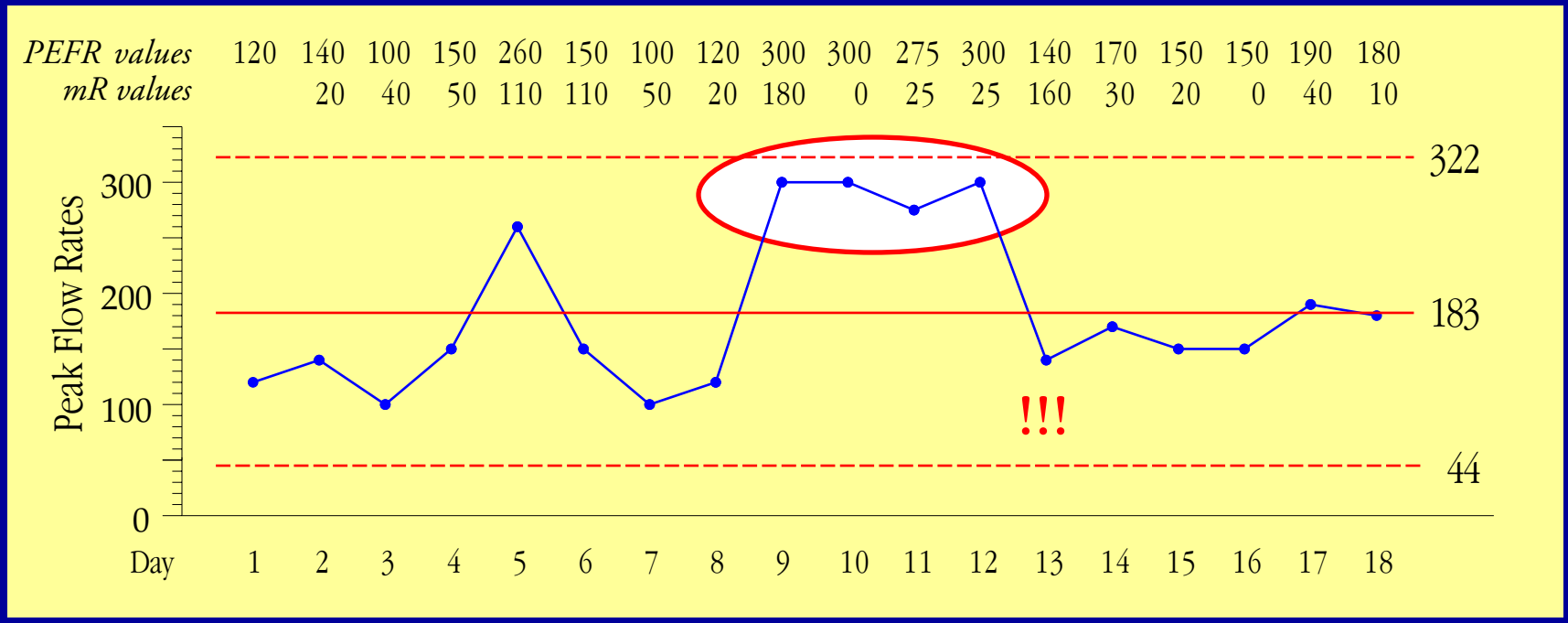
Why Susie Can Breathe

While standard allergy testing had shown that
this patient was allergic to dogs, dust mites, and common pollens,
and while these allergies had been explained
to both the patient and her mother,
this clinical knowledge did not have the impact
that the simple chart had.



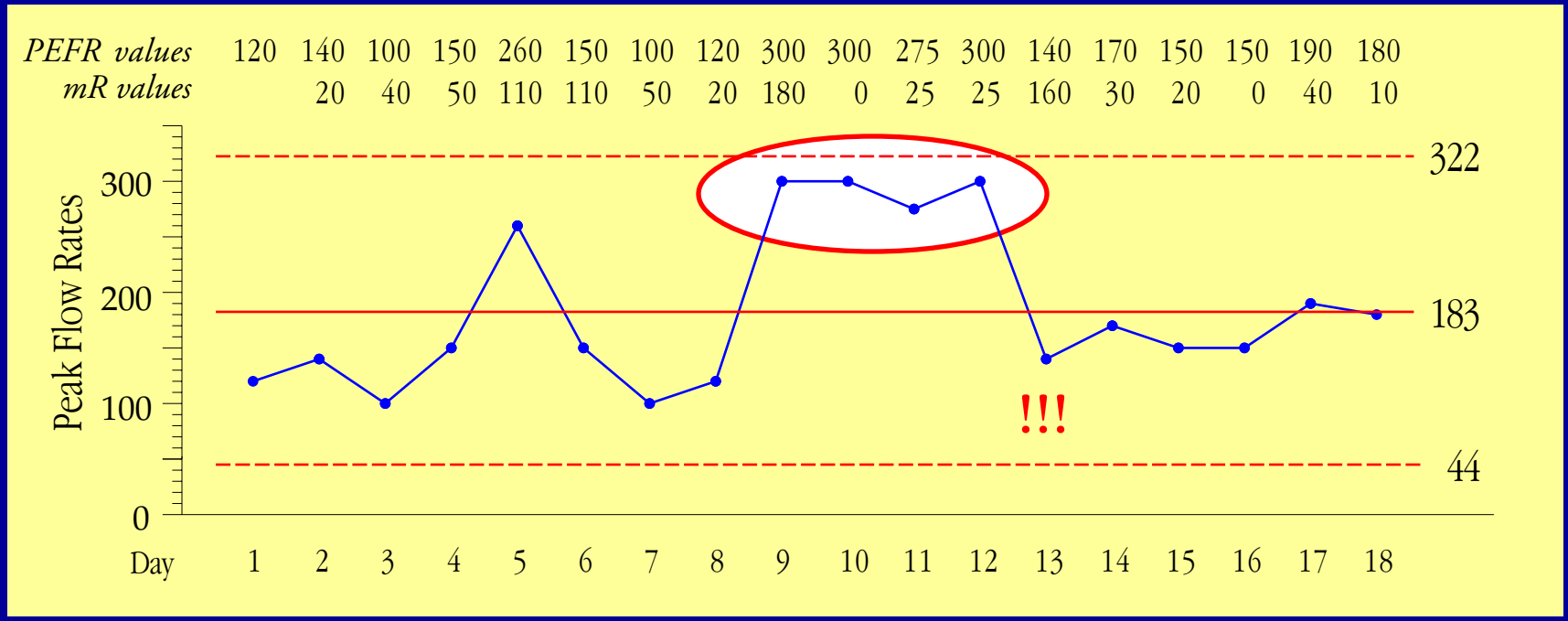
Why Susie Can Breathe

Having seen with their own data
the impact of these allergens in their own home
they began to make changes.
Both the patient and her mother changed
from being receivers and objects of treatment
to being active participants in that treatment.



Why Susie Can Breathe

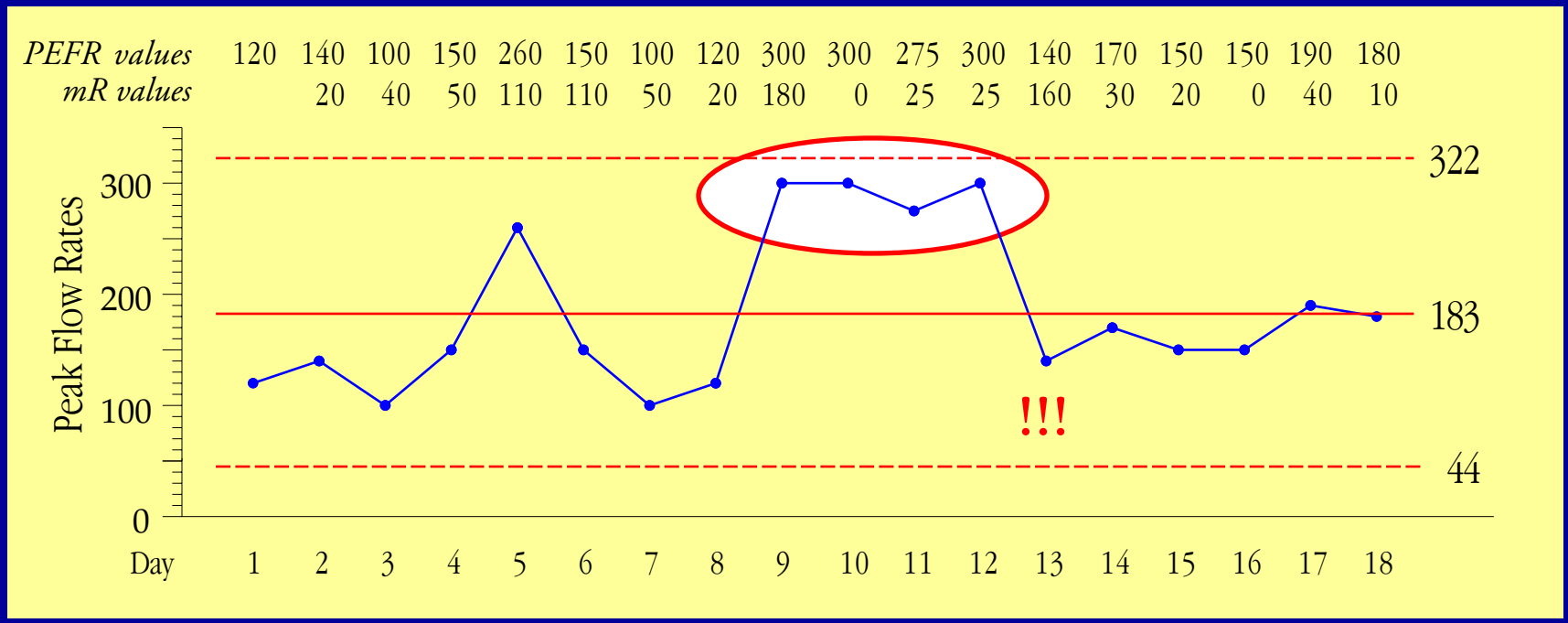
The most important part of any discovery
is the ability to communicate that discovery to others.



Process behavior charts are unequalled
at facilitating both discovery and communication.

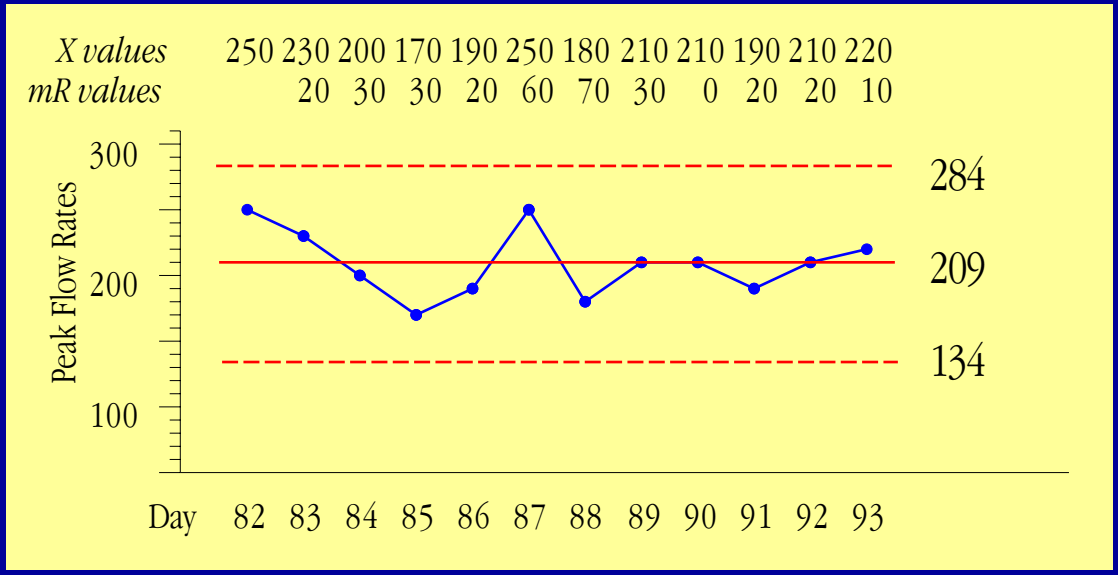
Why Susie Can Breathe

The chart also provided Dr. Boggs with an objective way to evaluate the effectiveness of treatment.
The 200 point swings from day to day are undesirable.
The lower limit of 44 is in the danger zone for severe asthma.
The average peak flow of 183 is only 60% of predicted best.
All of these need to change.



Why Susie Can Breathe

Based on the compelling evidence on the chart the mother and her daughter began to work to remove or lessen allergens and irritants in the patient's environment.

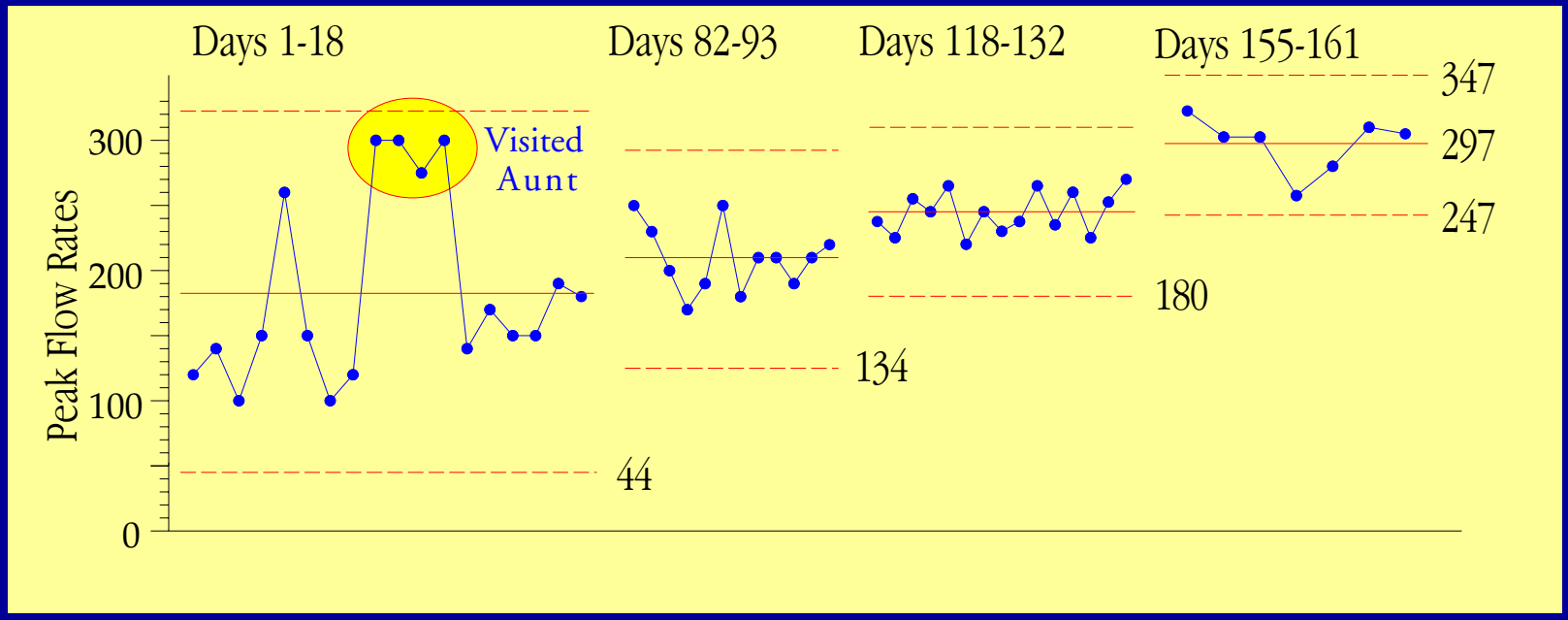


By days 82 to 93 they had made real progress.

Why Susie Can Breathe

Five months later the patient's average peak flow of 297
is 96% of her predicted best.
She has added 200 points to her lower limit
and she no longer suffers from severe asthma.

Now she is doing as well at home as she did at her aunt's,
and her bronchodilators can handle her asthma.

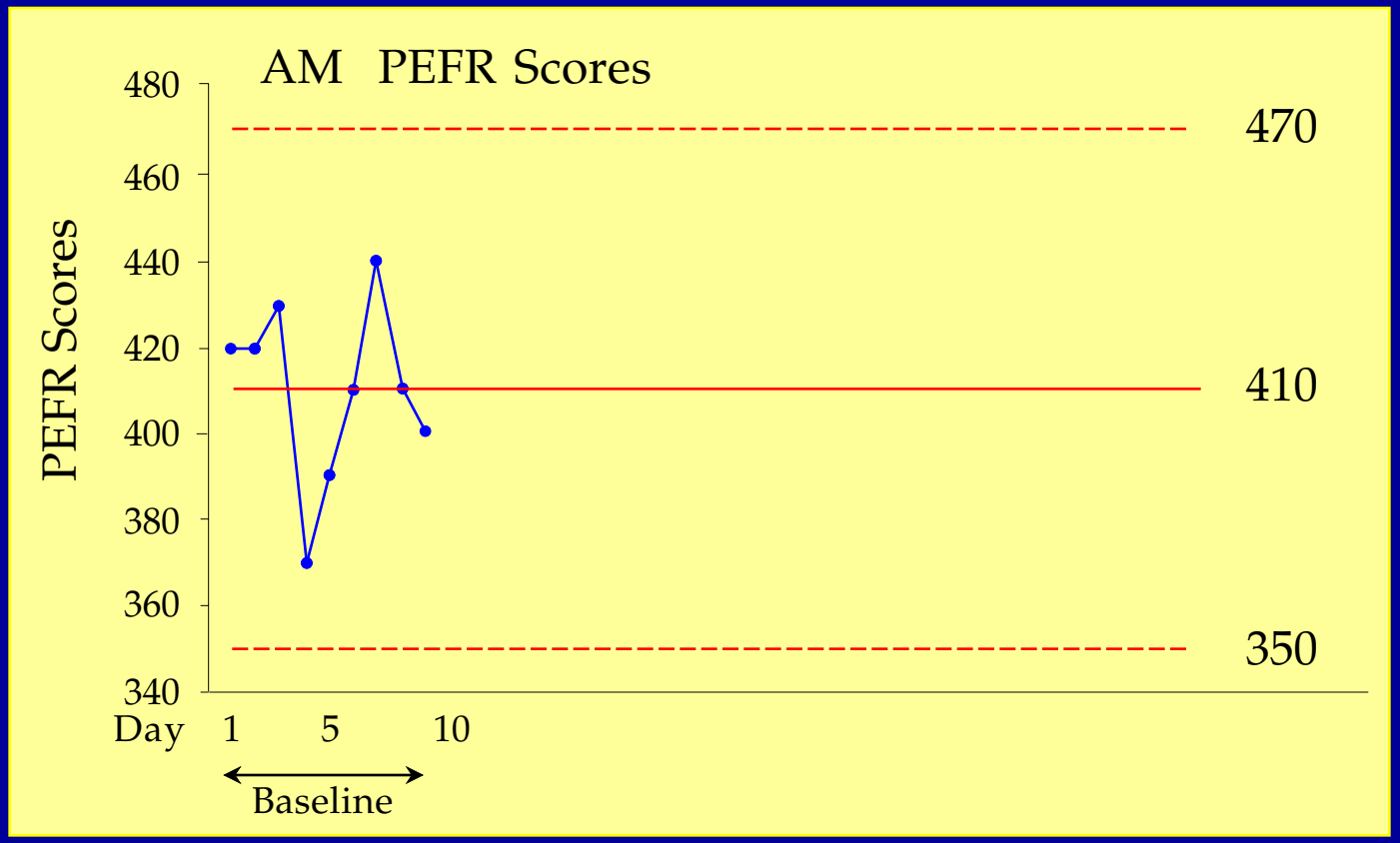


Jane Suffered from Rhinitis for 10 Years

A 37 year old had suffered from rhinitis for over 10 years.
Thinking it was allergies, she visited Dr. Boggs.
Skin tests showed no evidence
of allergen specific IgE antibodies.
So the rhinitis must be due to irritants.
As they searched for irritants in her environment,
Dr. Boggs decided to use PEFR scores
to monitor changes in her airways
even though she did not have asthma.

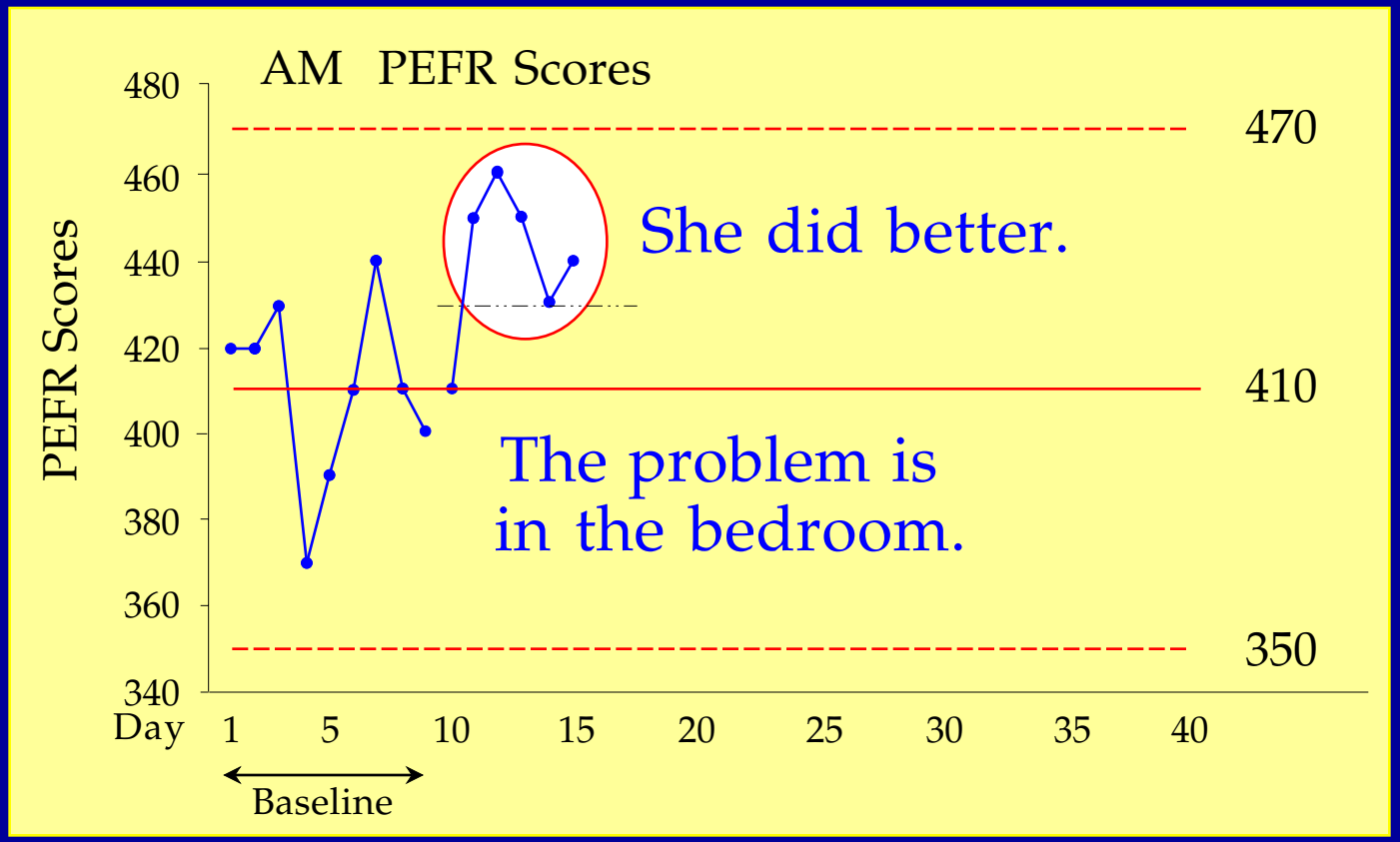
Jane Suffered from Rhinitis for 10 Years

They started with a 9-day baseline
where there were no deliberate changes
in the patient's home environment, work, or activities.



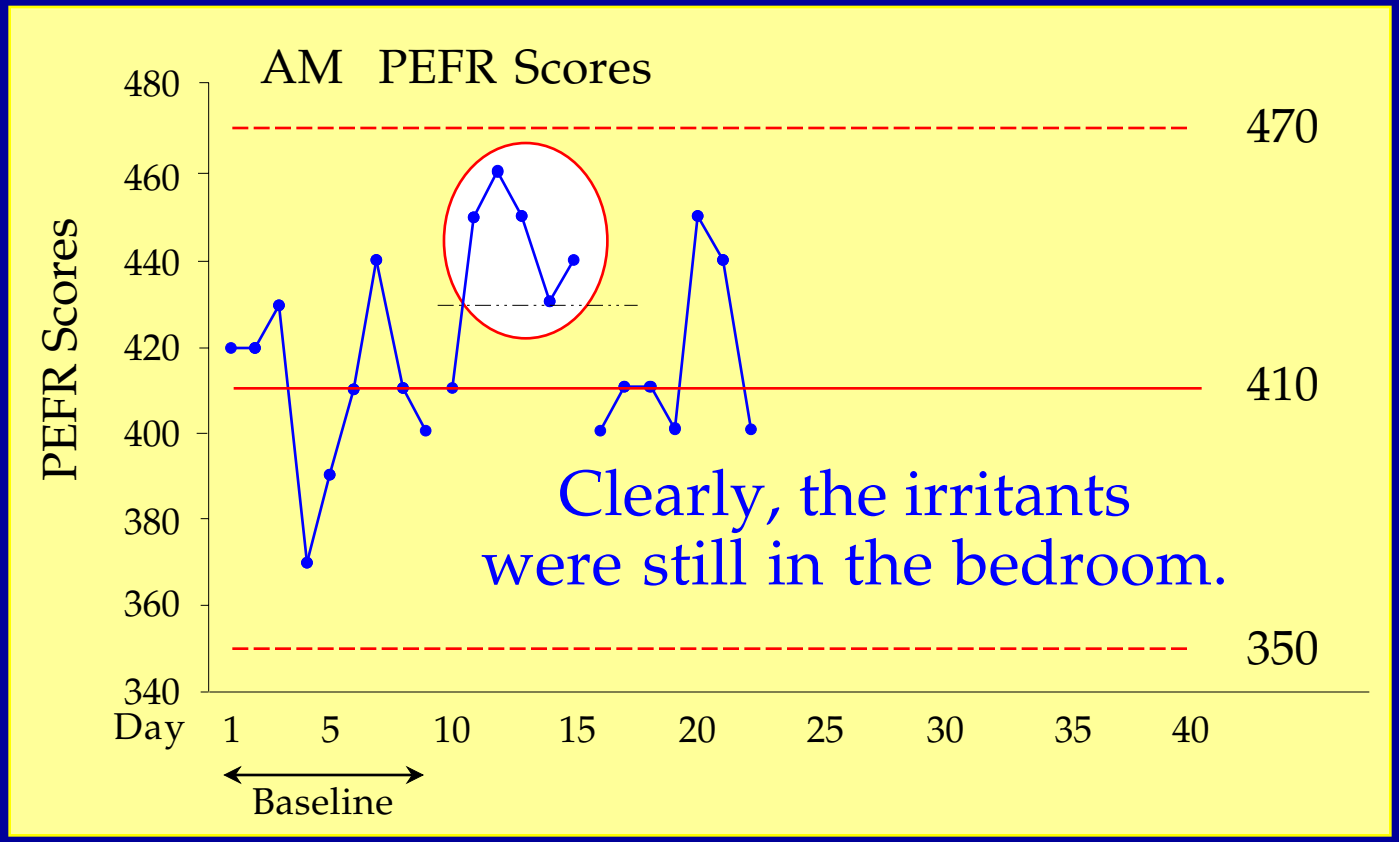
Jane Suffered from Rhinitis for 10 Years

Next the patient slept in the living room while they were changing the floor covering in the bedroom.

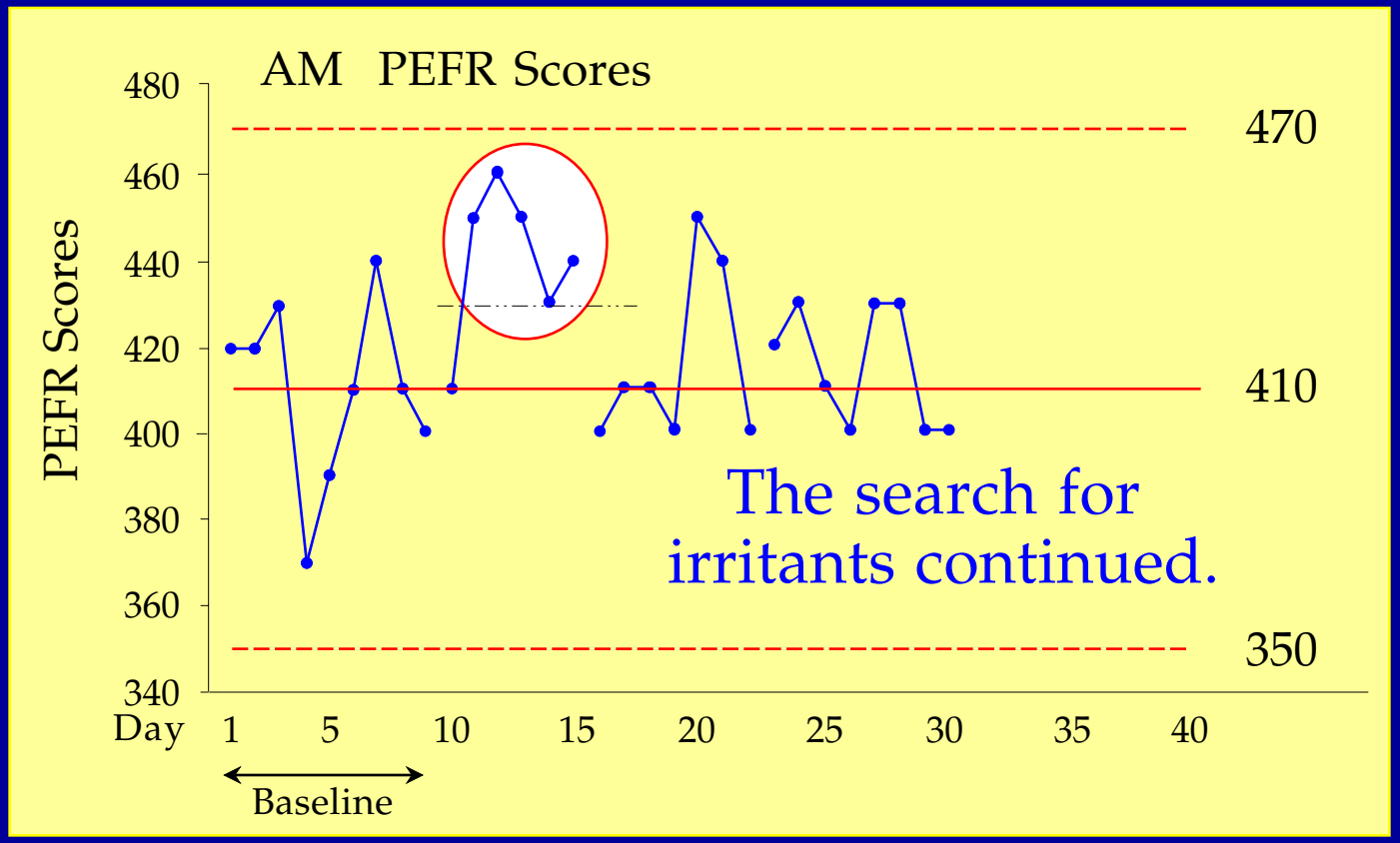


Jane Suffered from Rhinitis for 10 Years

When the patient moved back into the refurbished bedroom her rhinitis returned and her PEFR scores dropped back to baseline.



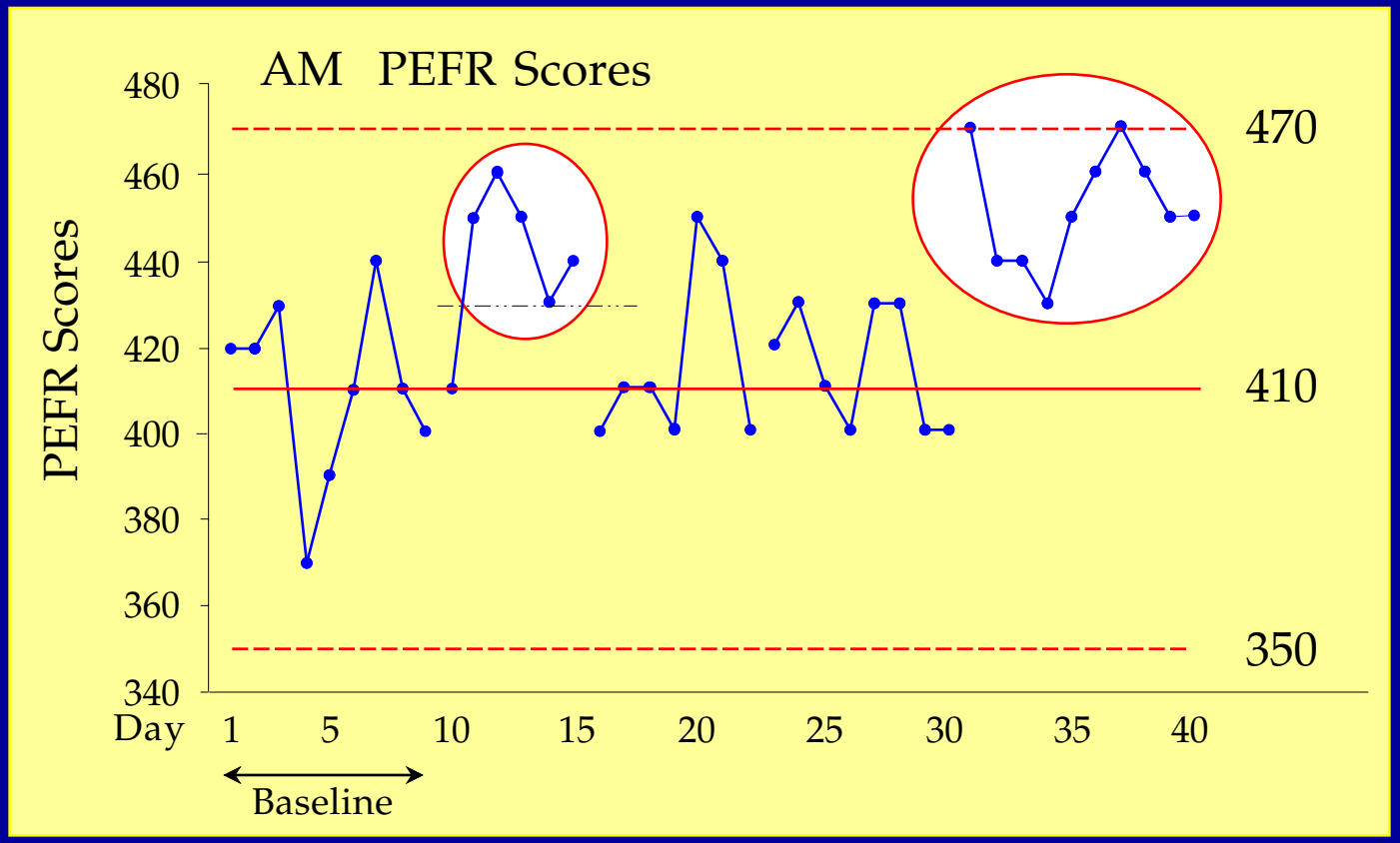
Jane Suffered from Rhinitis for 10 Years



They found that the attic access in the bedroom closet was completely open.

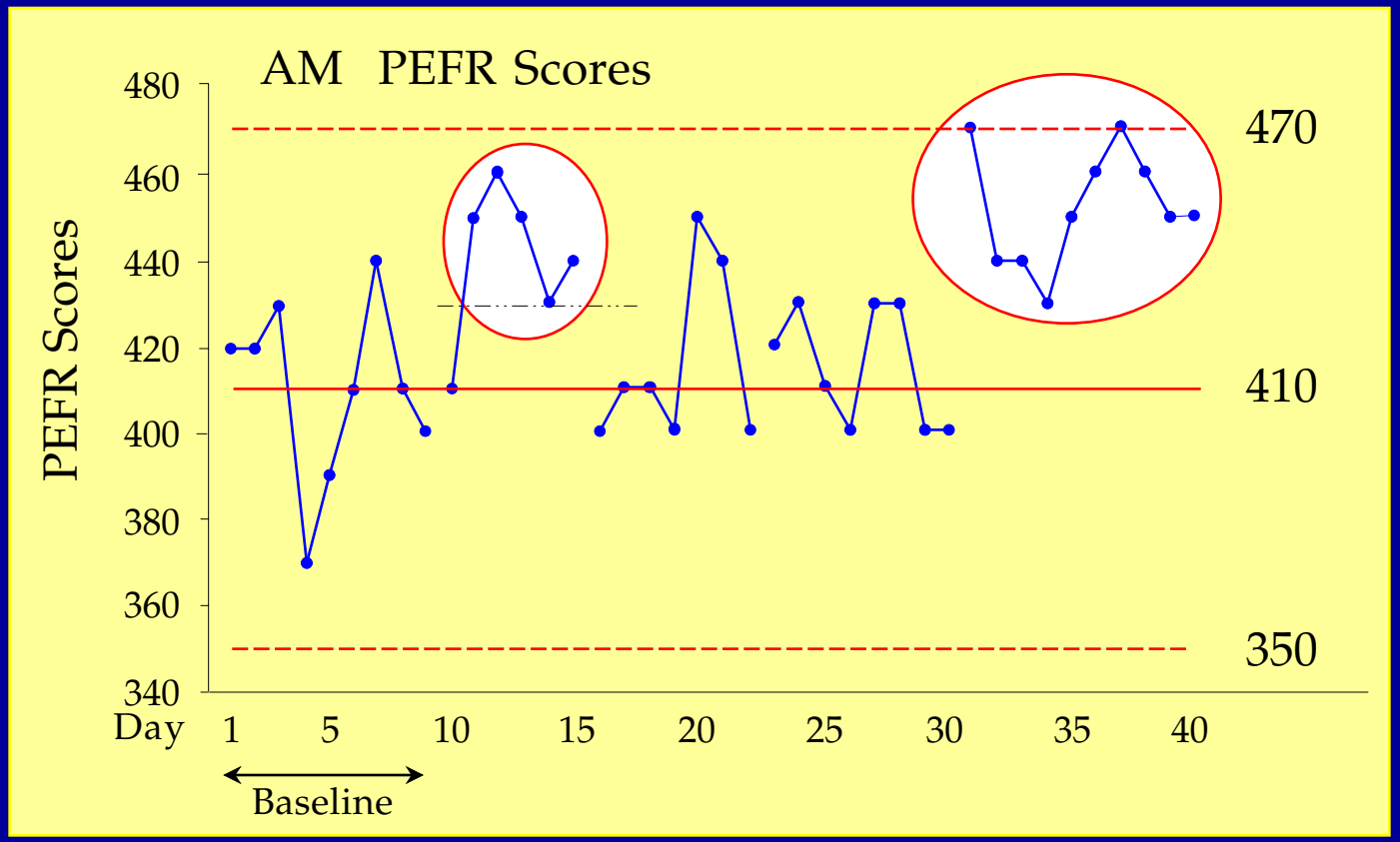
Jane Suffered from Rhinitis for 10 Years

When they sealed the hole in the closet ceiling there was no more insulation dust in the bedroom, and the patient had not felt this good in years!



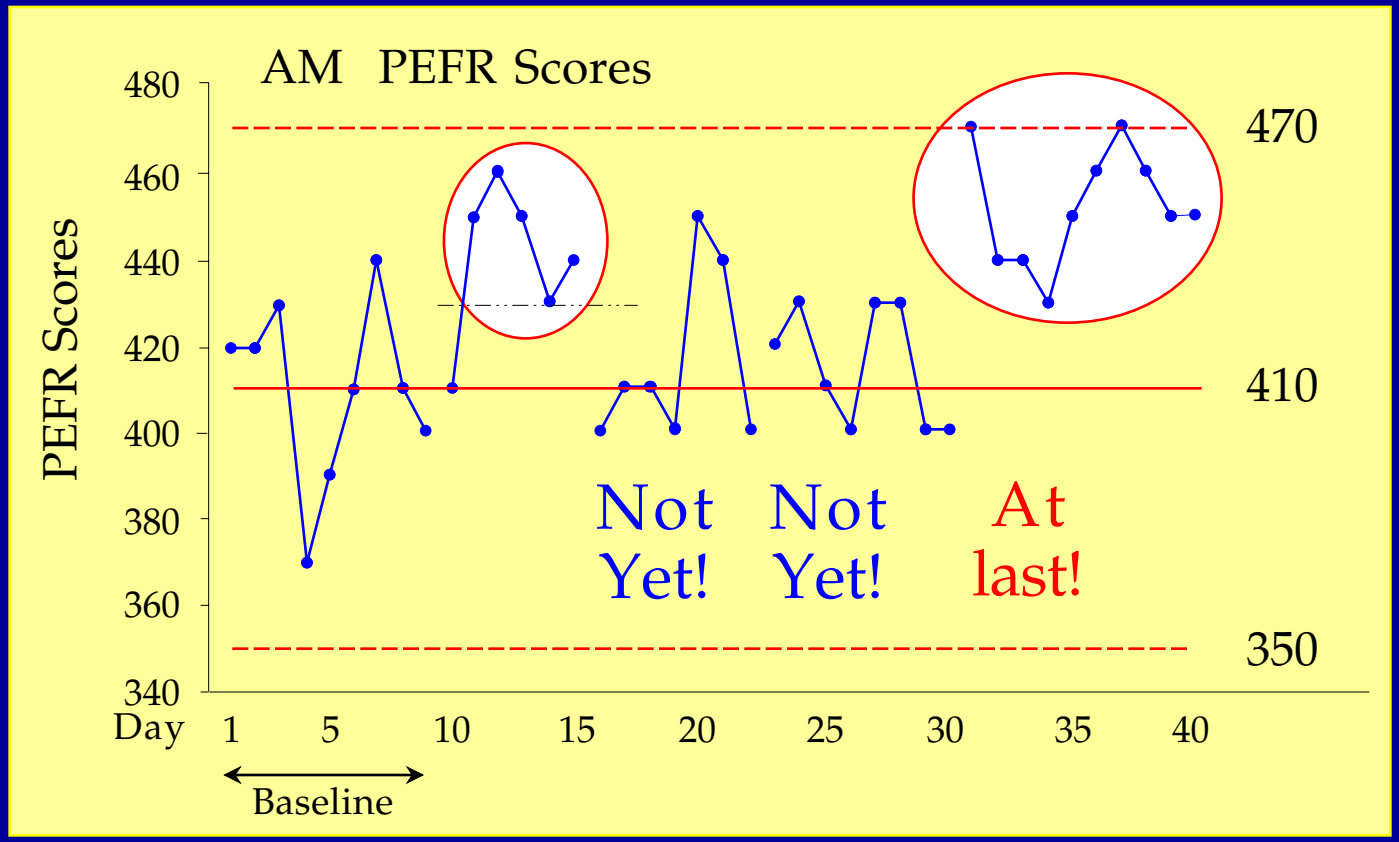
Jane Suffered from Rhinitis for 10 Years

Pity the husband if fixing the hole in the closet had been on his to-do list for the past 10 years!



Jane Suffered from Rhinitis for 10 Years

Process behavior charts provide
the feedback that is needed
to understand when our actions do not work,
and when they do work.



A Clinical Report: Peter's Story

Peter was a 25 year-old man with autistic disorder
who functioned in the severe range of mental retardation
and was blind from birth…

He commonly engaged in severe,
potentially life-threatening self-injurious behavior—
self-inflicted wounds caused by head-banging
and biting of the wrists and fingers

This is the story of incremental improvement
based upon the collection and analysis of data
in the context of the treatment program

A New Environment

Prior care had included
a variety of aversive behavioral treatment procedures,
including electric shocks…
these were all unsuccessful
in suppressing Peter's self-injurious behaviors.

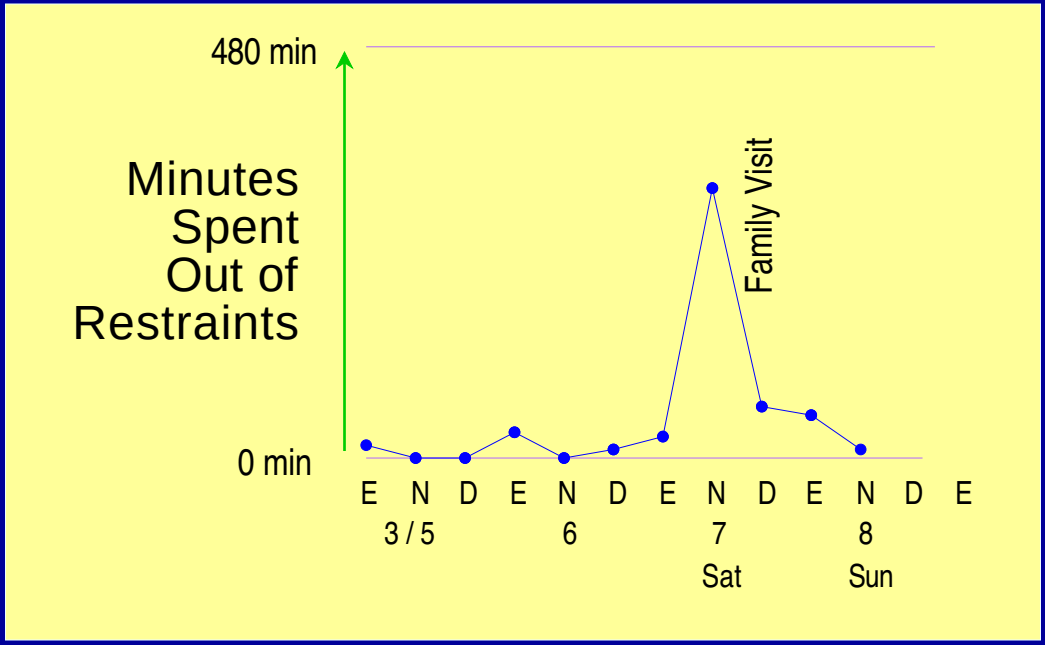
Our story starts as Peter was moved
into a new faclity

At the beginning of Peter's treatment
he had to be kept in restraints
which tethered his wrists to a belt
and he had to wear a football helmet
to prevent injury to his head.

A New Environment

The data recorded for Peter were

the total time spent out of restraints each shift.



These values could range from 0 minutes to 480 minutes

A New Environment

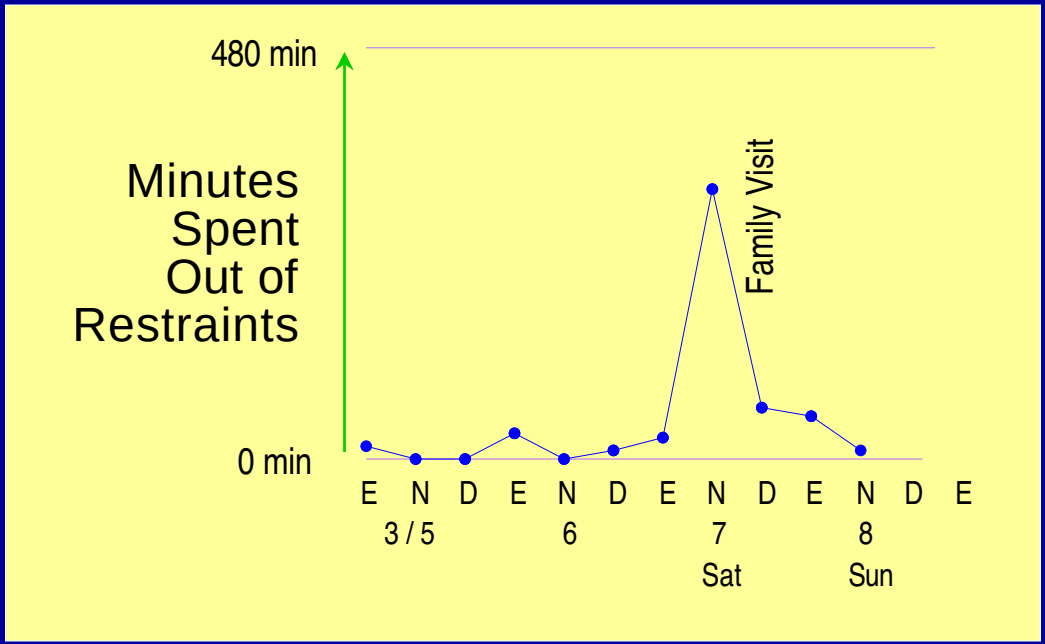
These values:

- reflected Peter's clinical status at the time of admission
- provided a focus for the treatment plan
- were objective and easy to obtain

  by using a simple timing device
- provided a record of the use of restraints

  as required by law

A New Environment



Initially, when they attempted to remove restraints,

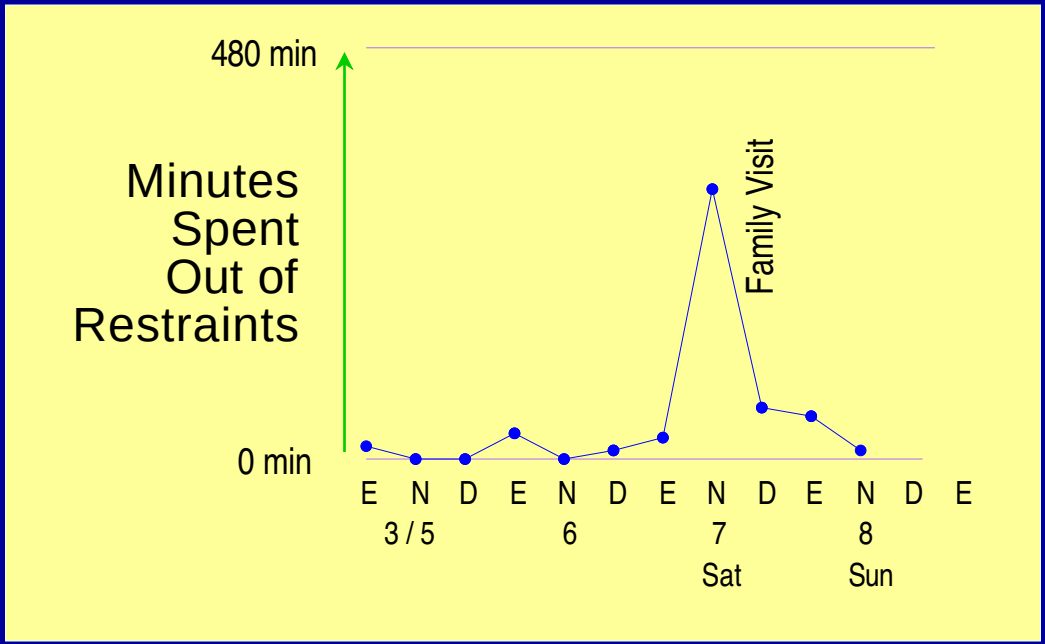
Peter became extremely agitated

and attempted to harm himself so violently

that he had to be immediately

returned to the restraints.

A New Environment



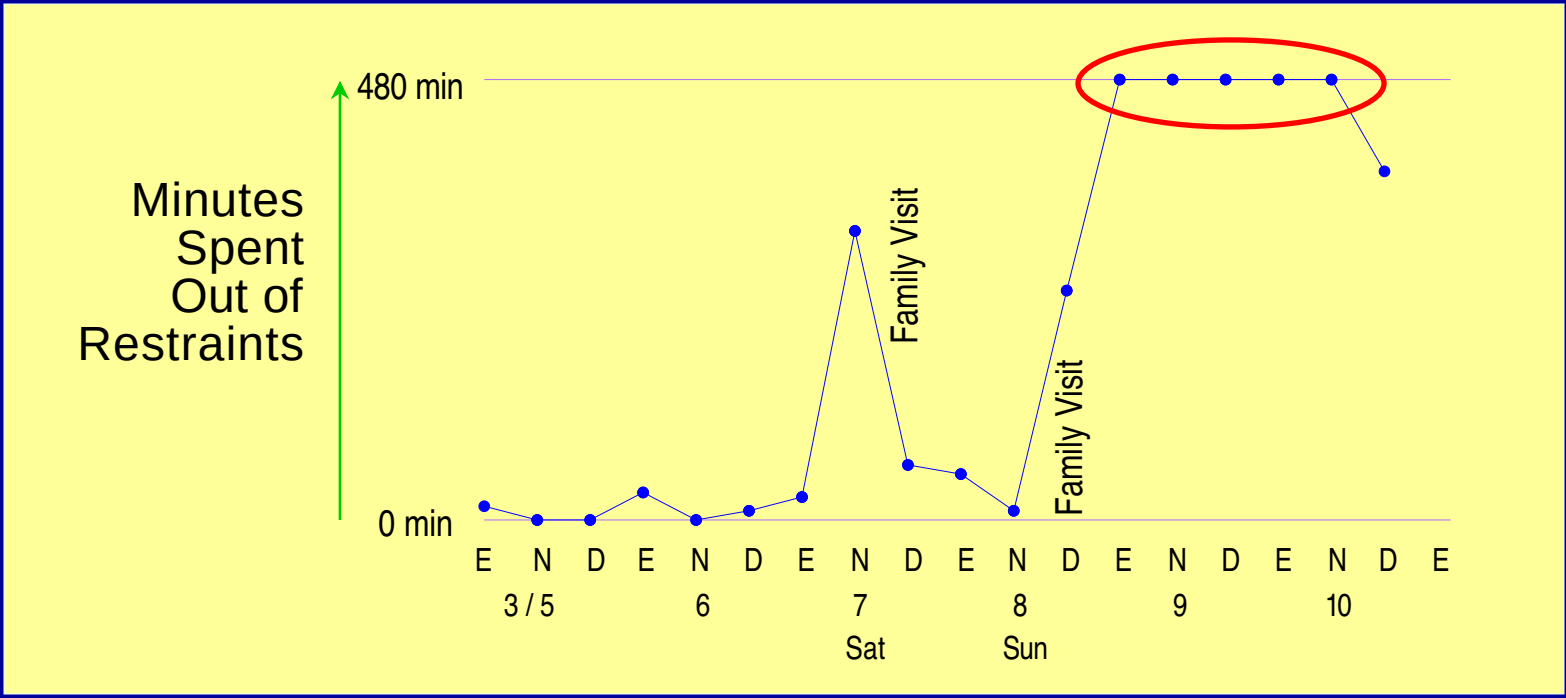
This suggested that the devices functioned
as a form of self-restraint
which is frequently seen among those with SIB

Thus, treatment would not attempt to prohibit restraints
but would focus on better ways
for Peter to ask for the restraints

A New Environment

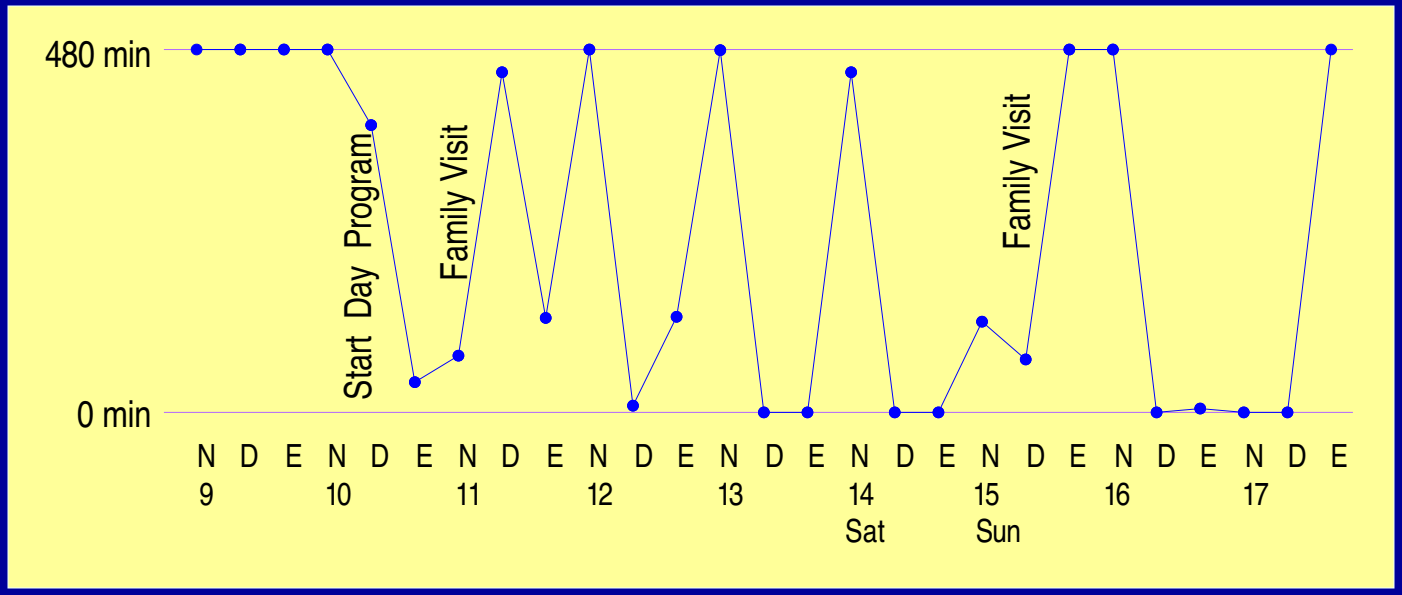
Effectiveness of this approach is seen in a few days—

Five straight shifts out-of-restraints…



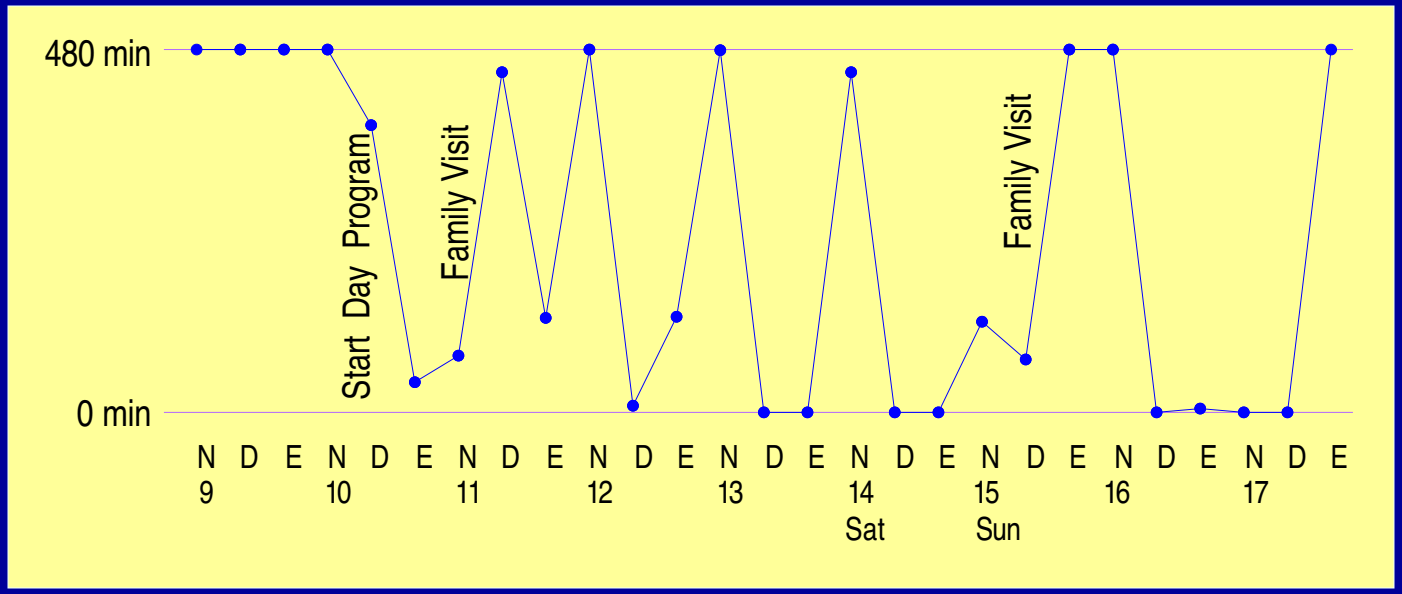
The Start of the Day Program

However, as level of task demands increased
at the start of the day program,
Peter reverted to the use of the restraints during day,
but continued to spend the nights without them.



The Start of the Day Program

The start of the day program is an understandably stressful experience for any blind person who has to learn how to navigate his way around a new environment, and even more so for autistic individual who may be agitated by changes in routine and exposure to strangers.

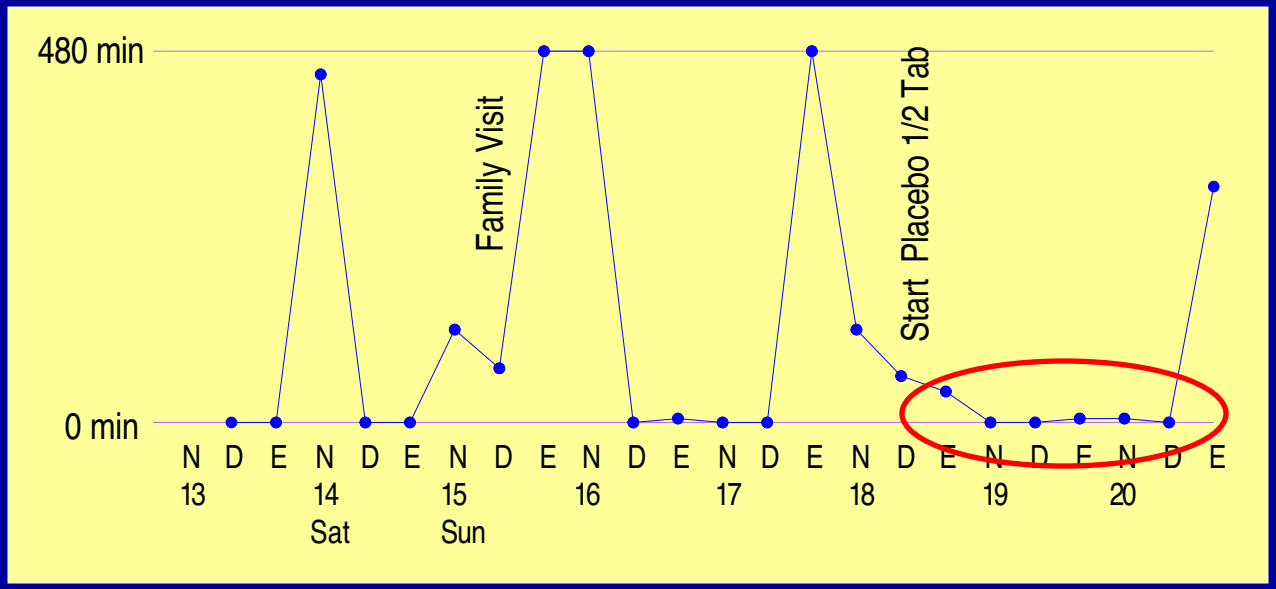


Treatment

Peter's next relapse

came at the start of the placebo—

another change in routine.



Treatment

When Peter entered the unit he was taking Drug A

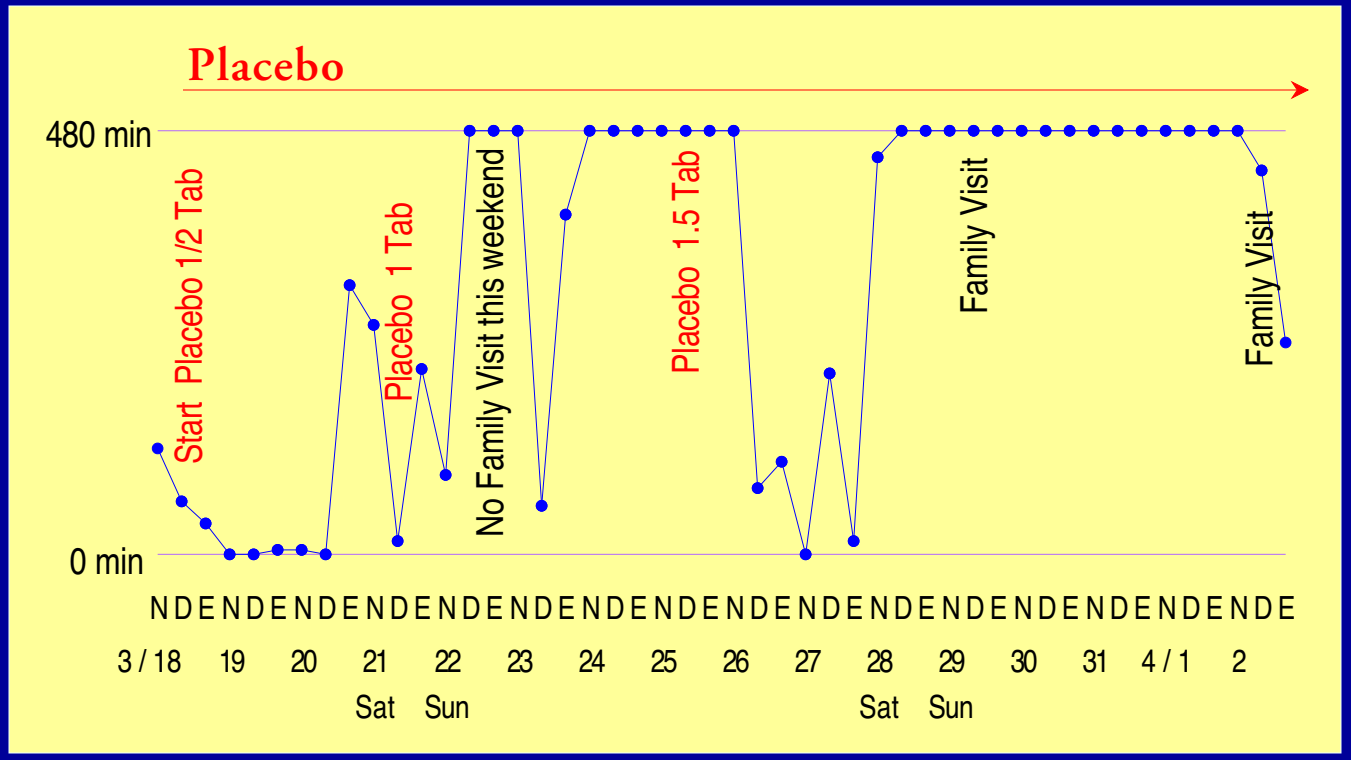
Treatment proposed the use of experimental Drug B in conjunction with Drug A

Therefore, a placebo treatment phase was planned to prepare for introduction of Drug B

During this phase Peter would get both Drug A and a Placebo

Treatment

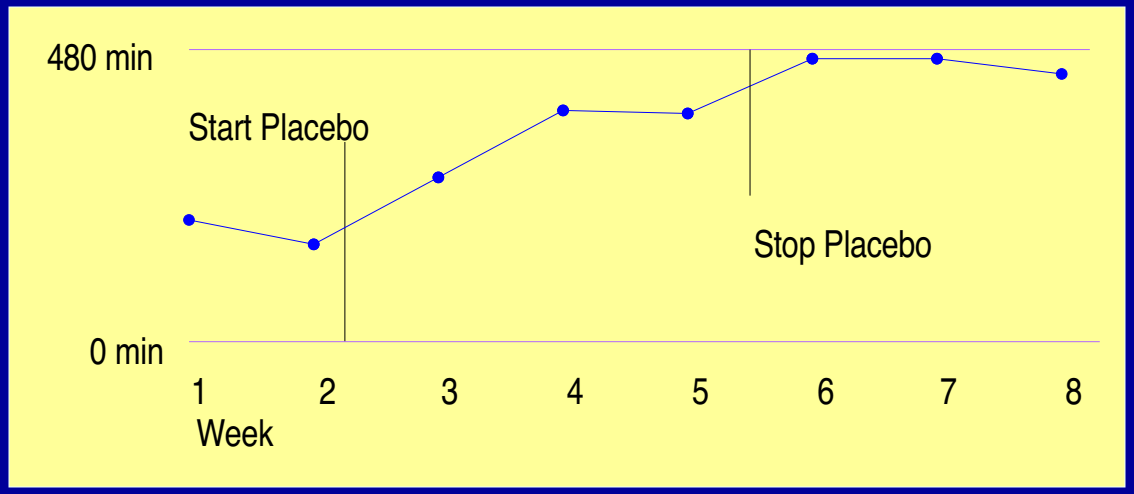
Both the staff impressions and the data suggested that Peter did better with the Placebo!



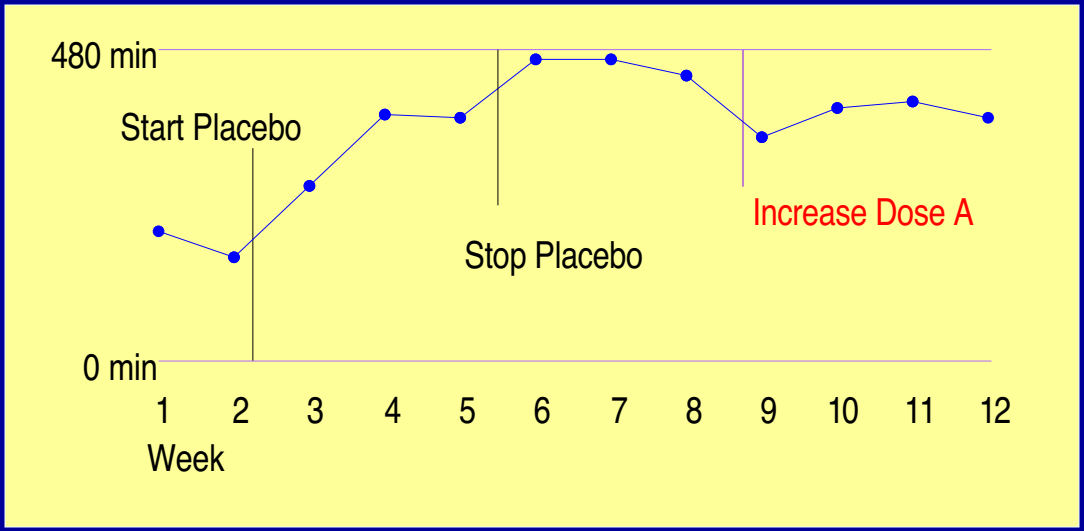
(It turned out that this "placebo" contained asprin)

Treatment

In evaluating overall changes in behavior
the psychologist felt that the shift-by-shift data
were a bit overwhelming,
and so he began to use weekly averages
to highligh the long-term trends



Treatment

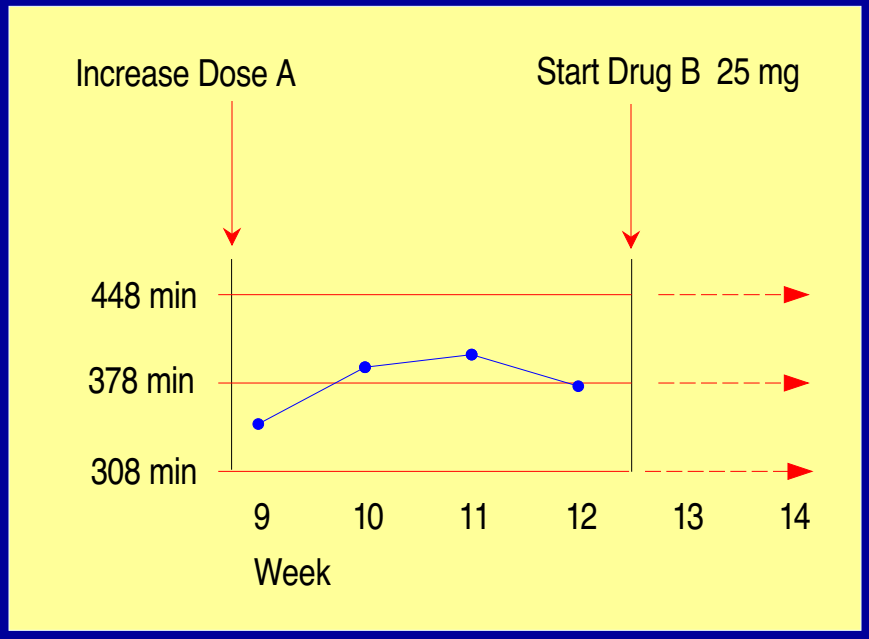


Prior to the introduction of Drug B, the dosage of Drug A was increased to evaluate possibility of helping Peter using only a single medication

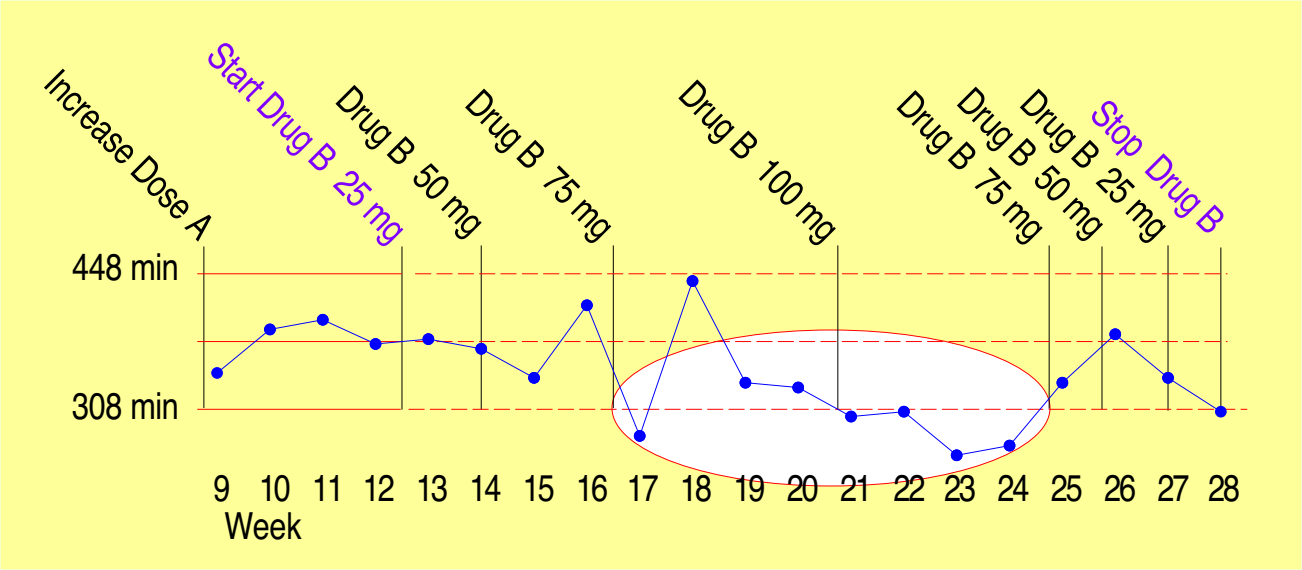
The chart above shows no positive effect from this increased level of Drug A

Treatment

The four weeks with increased levels of Drug A were taken as a baseline period and limits were computed from these data to use in evaluating the effect of Drug B.



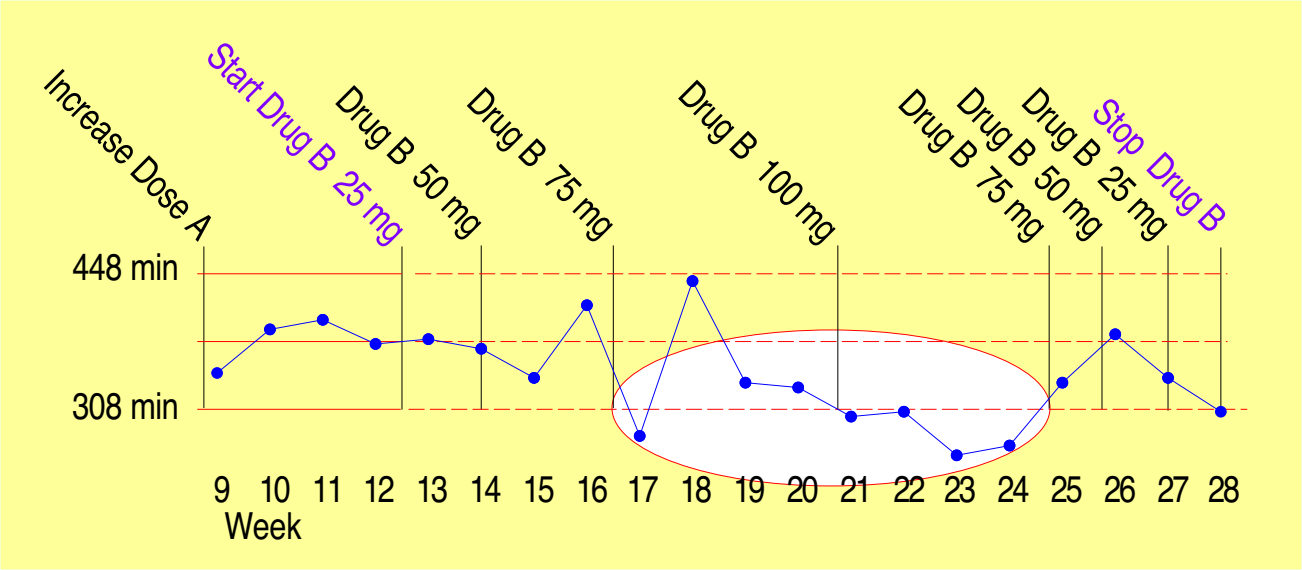
Treatment



As dosage of Drug B increased
the weekly averages dropped down
and Peter spent less time out of restraints

The higher dosages of Drug B corresponded
to a regression in Peter's clinical status

Treatment

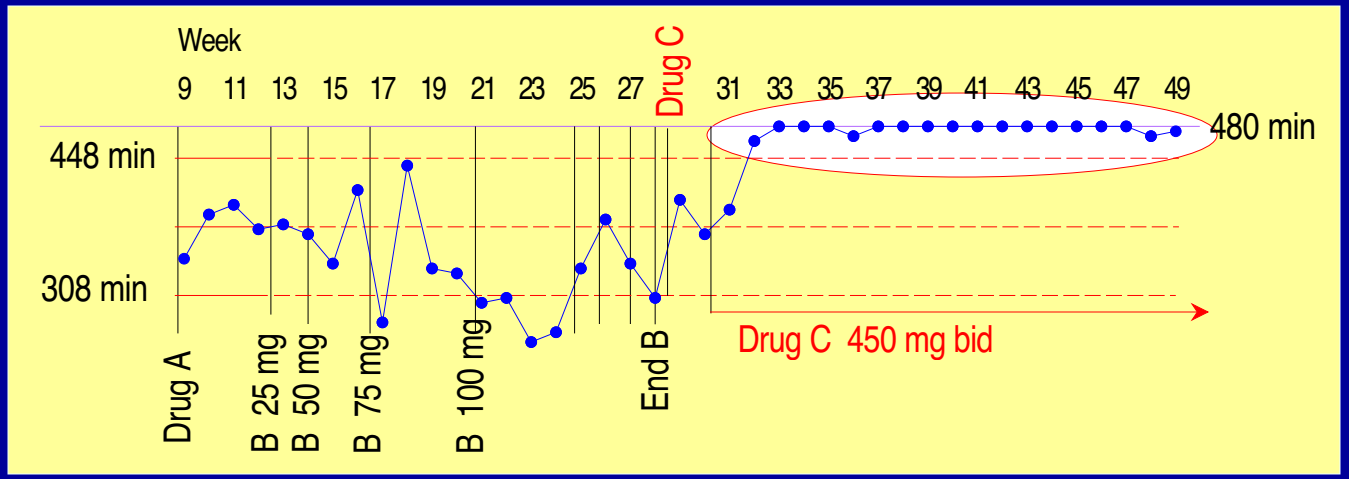


This regression was confirmed by points below lower limit as well as subjective impressions of staff—in fact Peter required sutures and antibiotics for facial injuries during this time.

When they reduced the dosage Peter spent more time out of restraints

Treatment

So Drug B was dropped and Drug C was introduced



Dramatic improvement in clinical status
both as seen on the chart and based
on subjective reports from the staff

Peter was essentially
completely free from the restriants
for the next 4 months

Feedback from Peter's Mother

"The statistical process of measuring our son's behavior has enabled those who have interpreted the graphs to have a clearer picture of what underlying factors influence the exhibited behaviors.

"Many behavior altering programs have been tried with varying responses in improved or worsening behavior. It has always been difficult, in a retrospective approach, to try to sort out those treatments which were effective. This sorting was usually left to subjective analysis and interpretive impressions.

Feedback from Peter's Mother

"We feel that this is a far more appropriate approach
toward the trials of medication
in attempts to control the self-injurious behavior.
In all the years of our son's behaviors, since the age of three,
we have never felt that a treatment program and staff
were as totally involved with him
as has been the case during the past year."

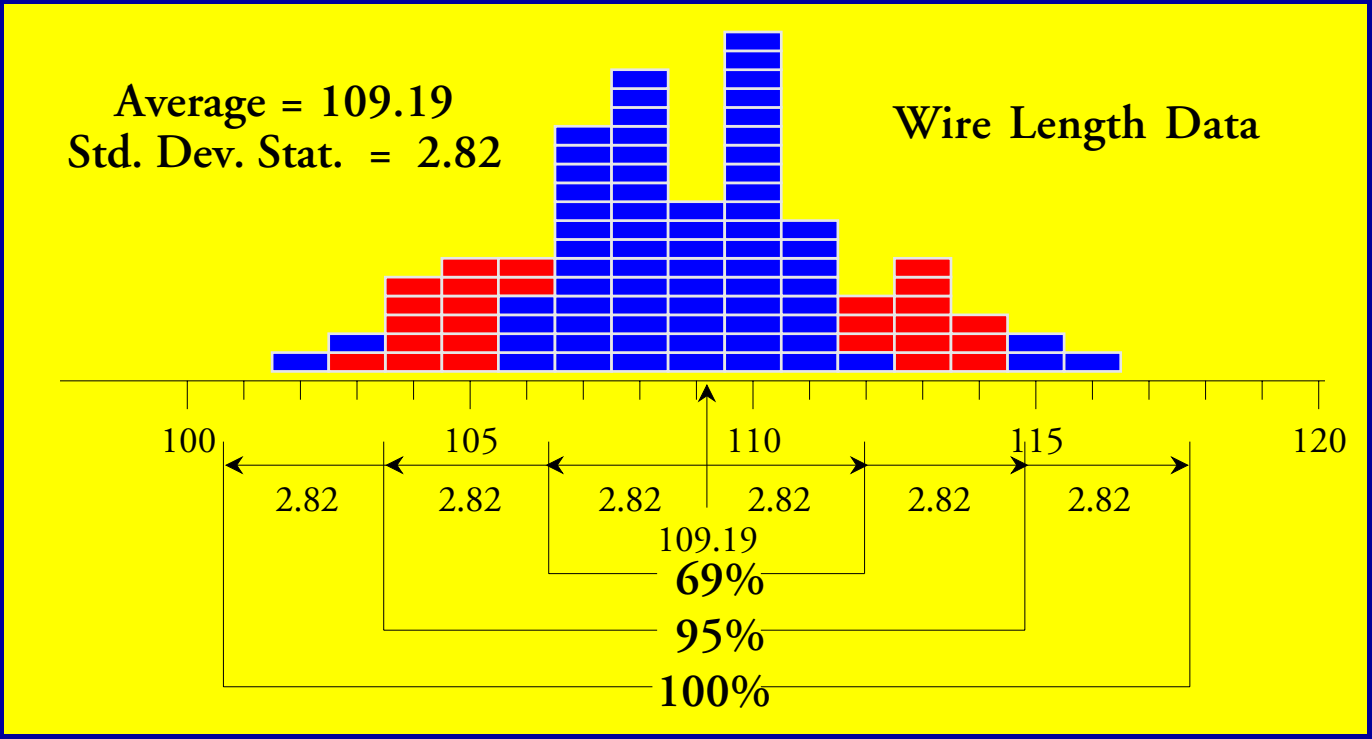
Summary

The visual display
of the available data
provides a clear and unequivocal way
of understanding and evaluating
both the status of the client
and the apparent effectiveness of the treatment.

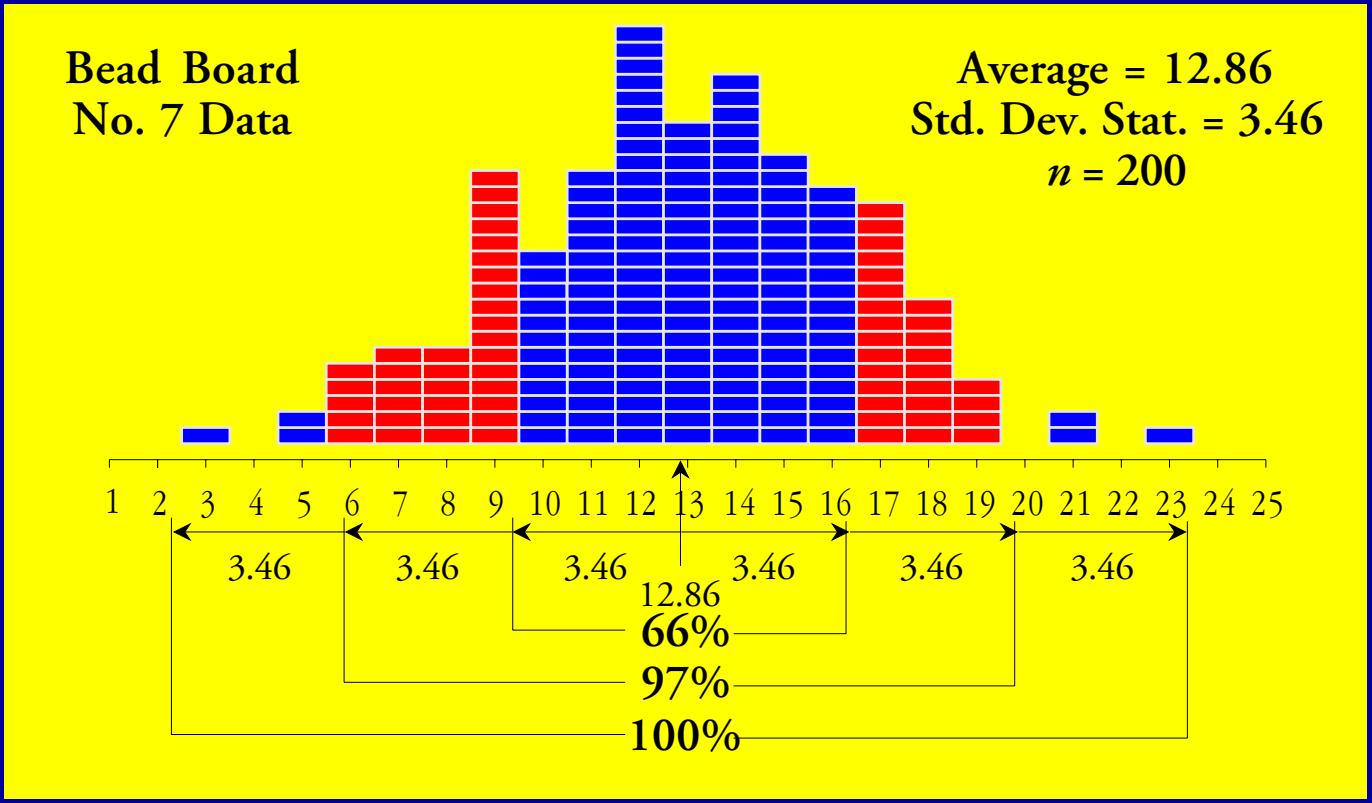
The charts are the voice of the process,
and this voice is an important part
of any treatment strategy.

WHAT DESCRIPTIVE STATISTICS DO

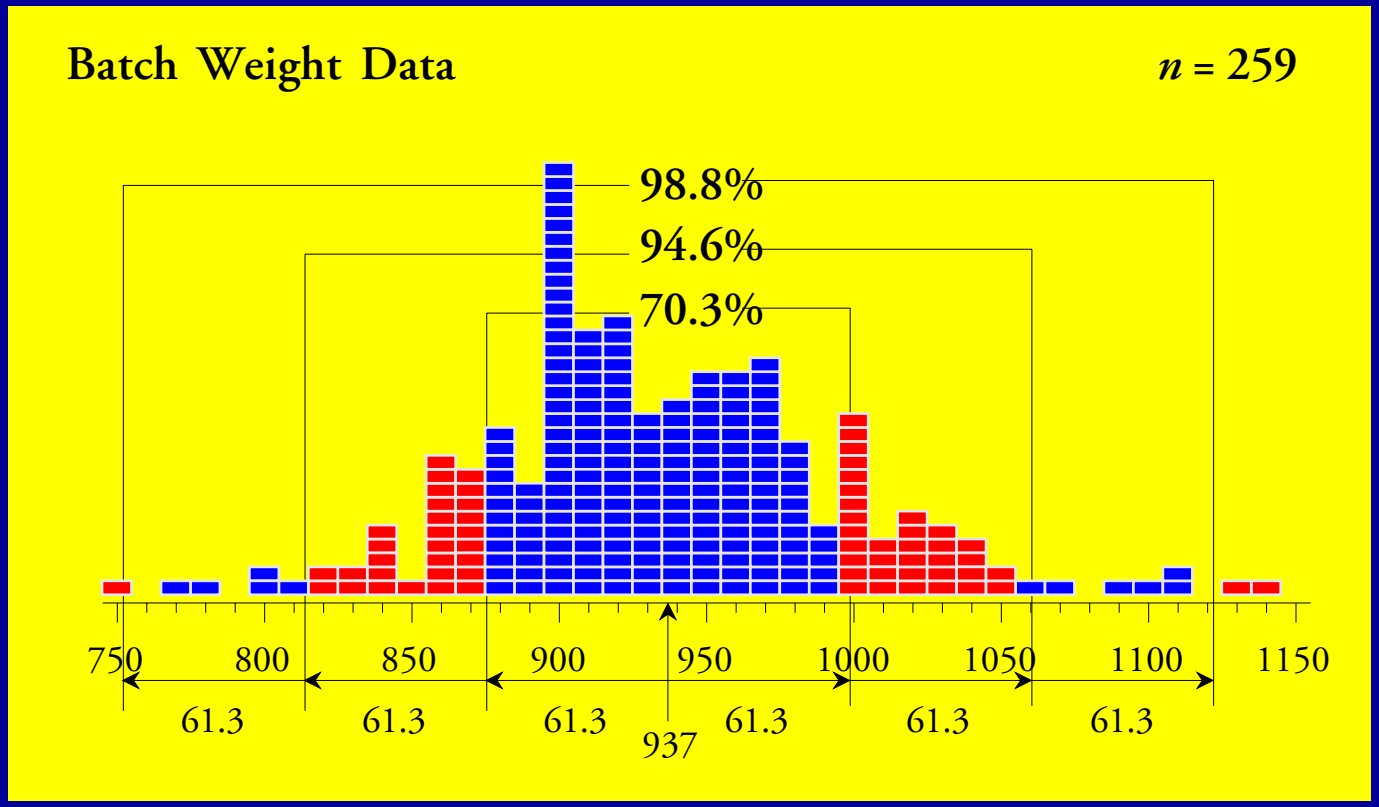
Descriptive Statistics Characterize the Data:
The Average characterizes Location, and
the Standard Deviation Statistic characterizes Dispersion.
Nested Intervals centered on the Average,
with radius equal to the Standard Deviation Statistic
will bracket certain proportions of the data.



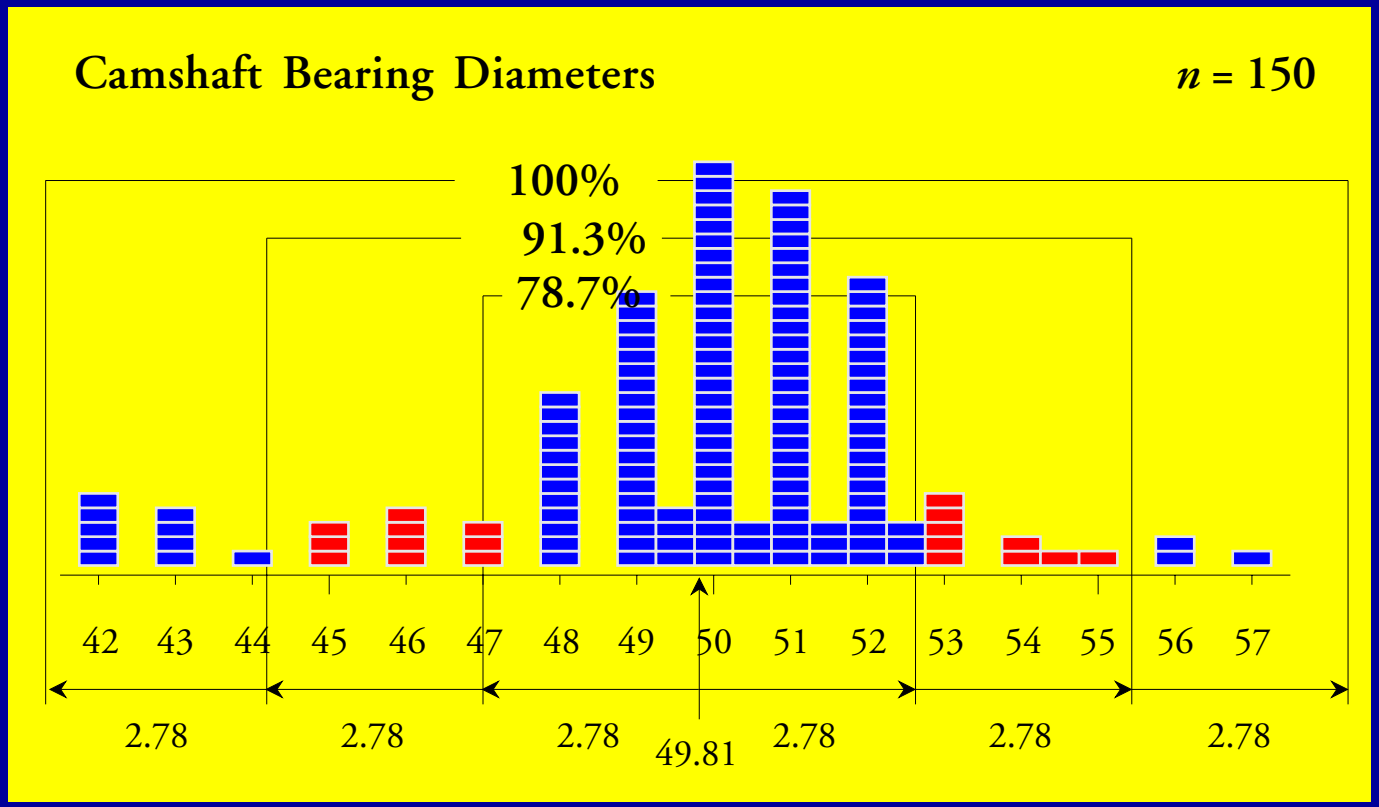
WHAT DESCRIPTIVE STATISTICS DO



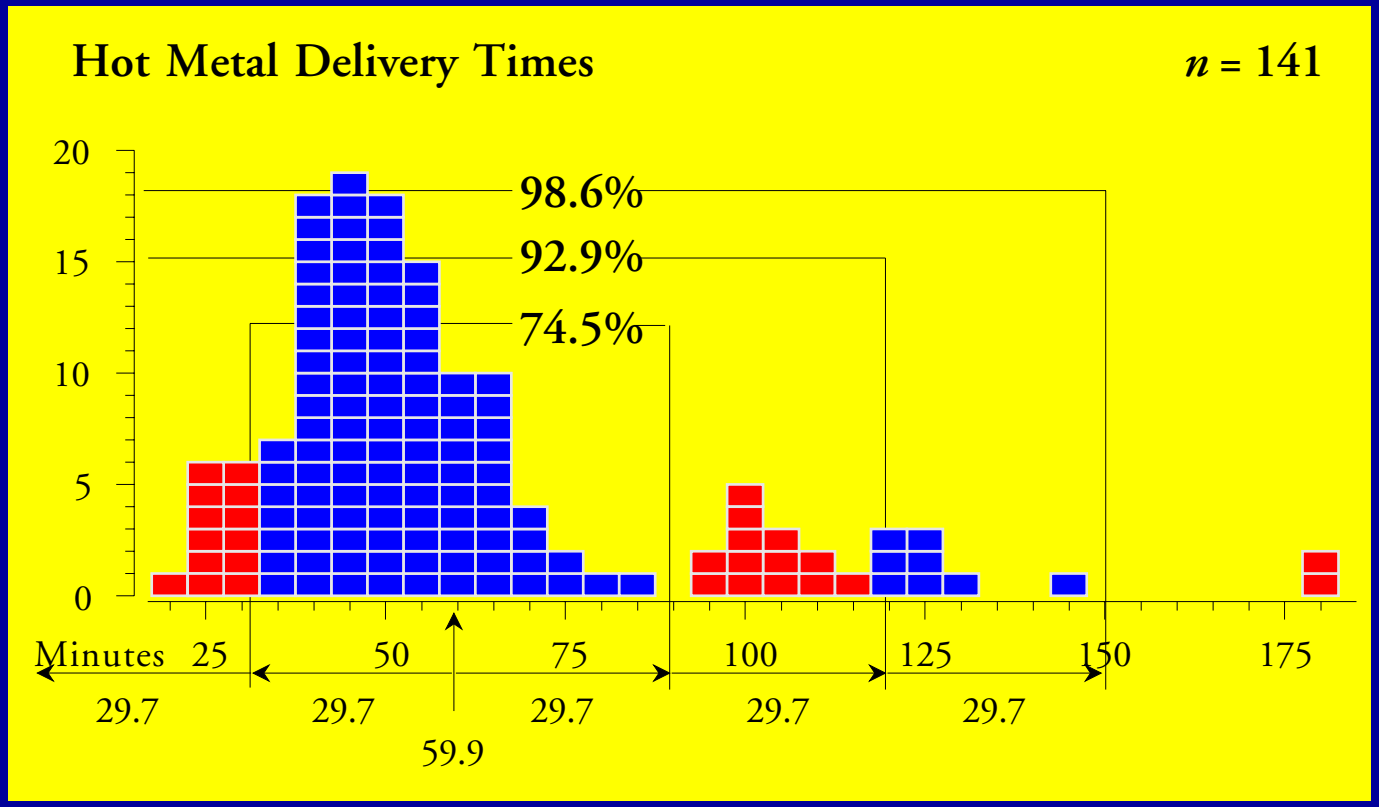
WHAT DESCRIPTIVE STATISTICS DO



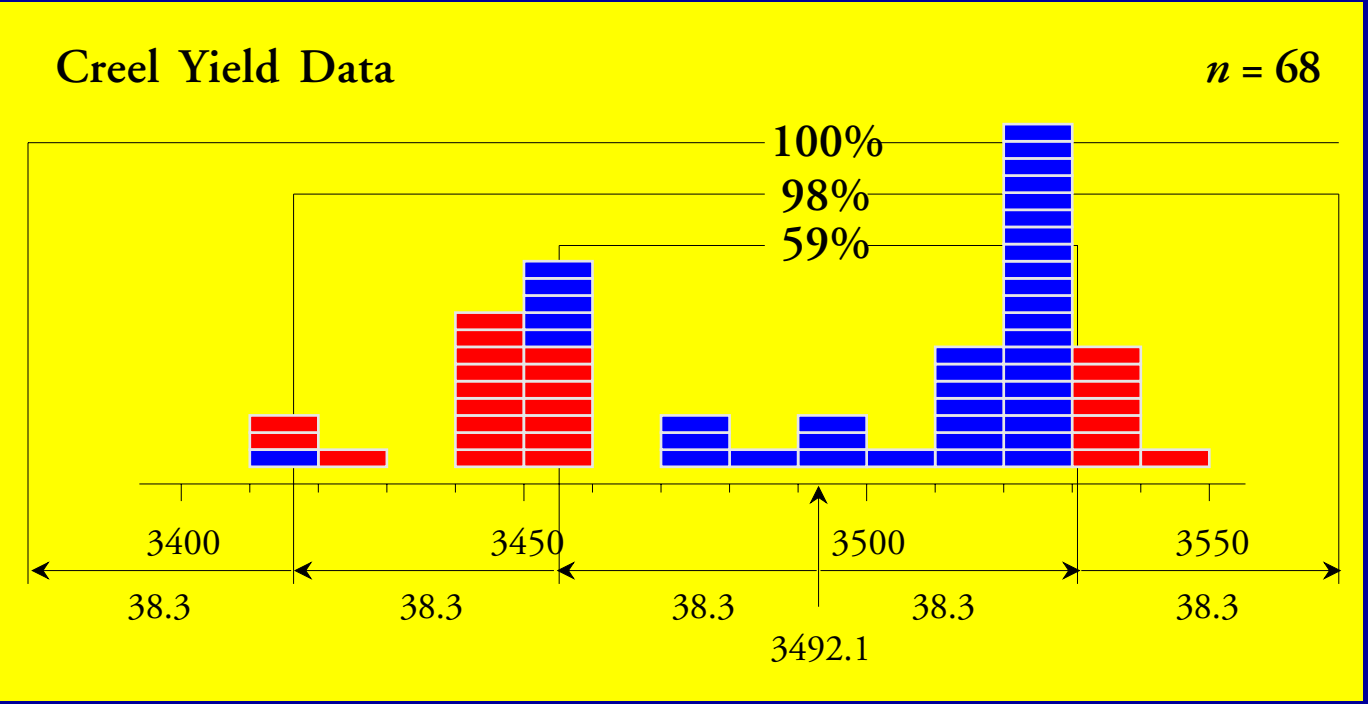
WHAT DESCRIPTIVE STATISTICS DO



WHAT DESCRIPTIVE STATISTICS DO



WHAT DESCRIPTIVE STATISTICS DO



WHAT DESCRIPTIVE STATISTICS DO

Descriptive Statistics Characterize the Data:

	Average ± 1 Std. Dev.	Average ± 2 Std. Dev.	Average ± 3 Std. Dev.
Wire Length Data	69%	95%	100%
Bead Board No. 7 Data	66%	97%	100%
Batch Weight Data	70.3%	94.6%	98.8%
Bearing Diameters	78.7%	91.3%	100%
Hot Metal Deliveries	74.5%	92.9%	98.6%
Creel Yield Data	59%	98%	100%

Average ± Standard Deviation Statistic will bracket the bulk of the data.

Average ± 2 Standard Deviation Statistics will bracket at least 90% of the data.

Average ± 3 Standard Deviation Statistics will bracket virtually all of the data.

WHAT DESCRIPTIVE STATISTICS DO NOT DO

Descriptive Statistics summarize the data.
They do not characterize
the Underlying Process which generated the Data.
They assume the data are homogeneous,
they do not check to see if this is so.
Some of our data sets are nonhomogeneous.

	Average ± 1 Std. Dev.	Average ± 2 Std. Dev.	Average ± 3 Std. Dev.
Wire Length Data	69%	95%	100%
Bead Board No. 7 Data	66%	97%	100%
Batch Weight Data	70.3%	94.6%	98.8%
Bearing Diameters	78.7%	91.3%	100%
Hot Metal Deliveries	74.5%	92.9%	98.6%
Creel Yield Data	59%	98%	100%

Three Standard Deviation Limits do not allow you
to reliably separate homogeneous data sets
from nonhomogeneous data sets.

WHAT DESCRIPTIVE STATISTICS DO NOT DO

Descriptive Measures of Dispersion do not differentiate between homogeneous data sets and nonhomogeneous data sets because they are computed globally.

If data are not homogeneous, the global measures of disperison will be inflated by the lack of homogeneity, however, localized measures of dispersion will not be inflated in the same way, or to the same extent, as the global measures.

Thus, to examine the data for possible evidence of nonhomogeneity, and to thereby characterize the Underlying Process we MUST use Localized Measures of Dispersion.

LOCALIZED DISPERSION

When the data have a time-ordering,
we arrange the data in time order,
and find the differences between successive values.

These differences summarize the local, short-term variation
and are called moving ranges.

We then summarize the moving ranges by computing
the Average Moving Range or the Median Moving Range.

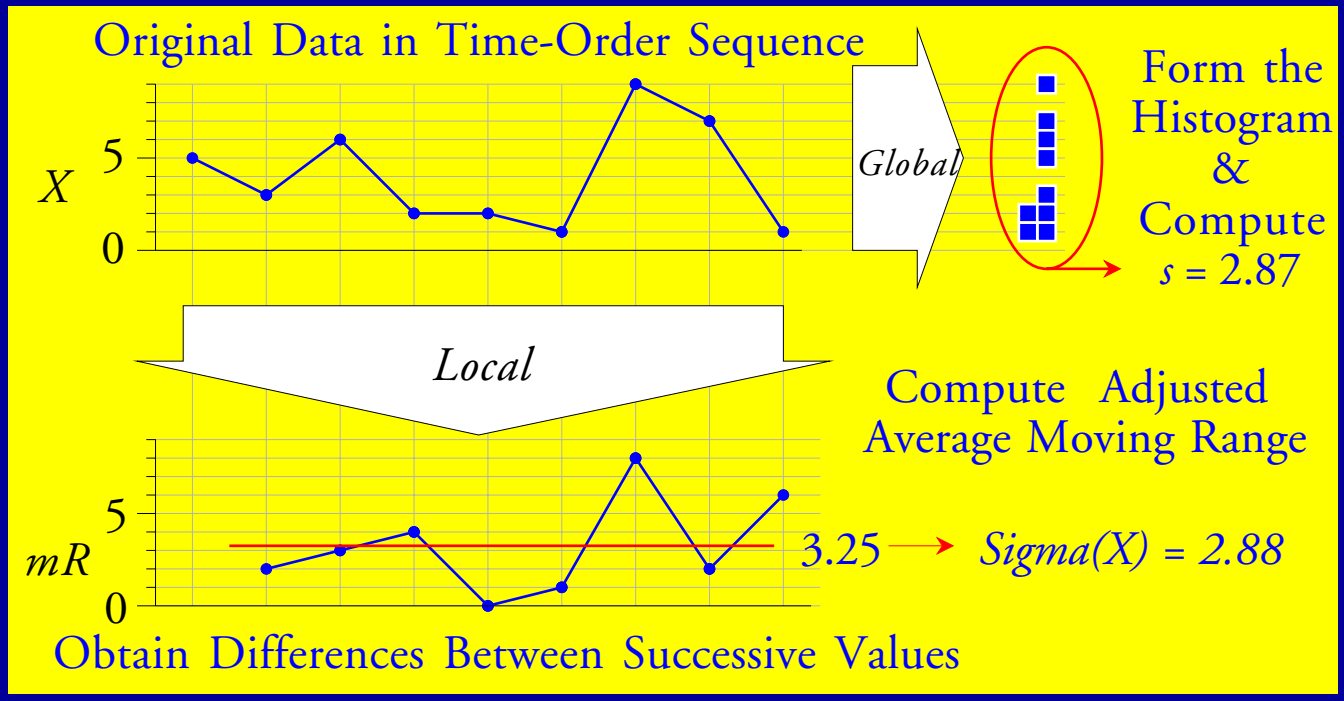
When divided by the appropriate constant,
these summary statistics provide a Local Measure of Dispersion:

$$\textit{Sigma}(X) = \text{Average Moving Range} / 1.128$$

or

$$\textit{Sigma}(X) = \text{Median Moving Range} / 0.954$$

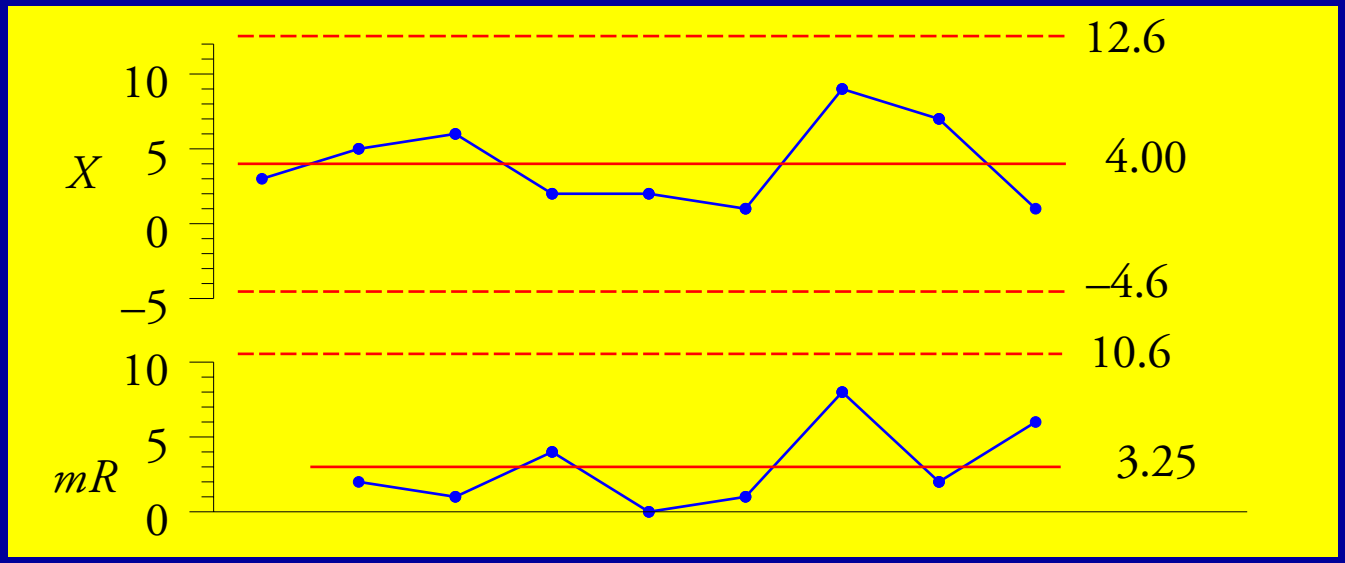
LOCALIZED DISPERSION



When the data are homogeneous
the global measure of disperison
and the local measure of dispersion will be quite similar.

When the data are not homogeneous
the global measure of dispersion
will be inflated relative to the local measure of dispersion.

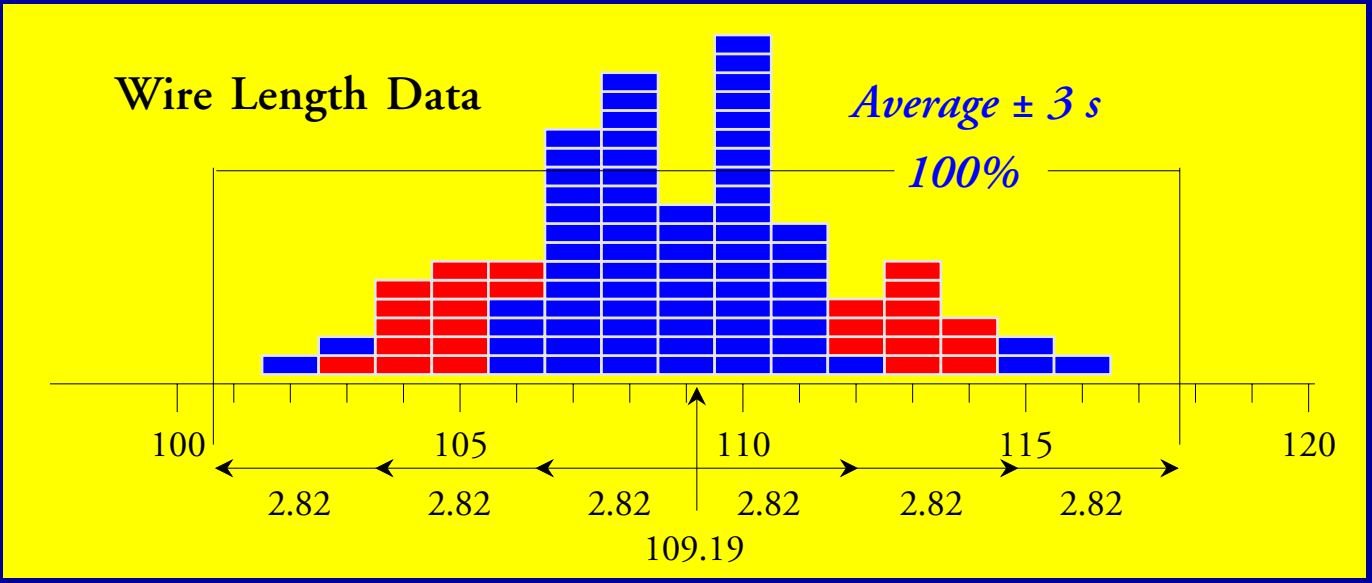
THE HOMOGENEITY QUESTION

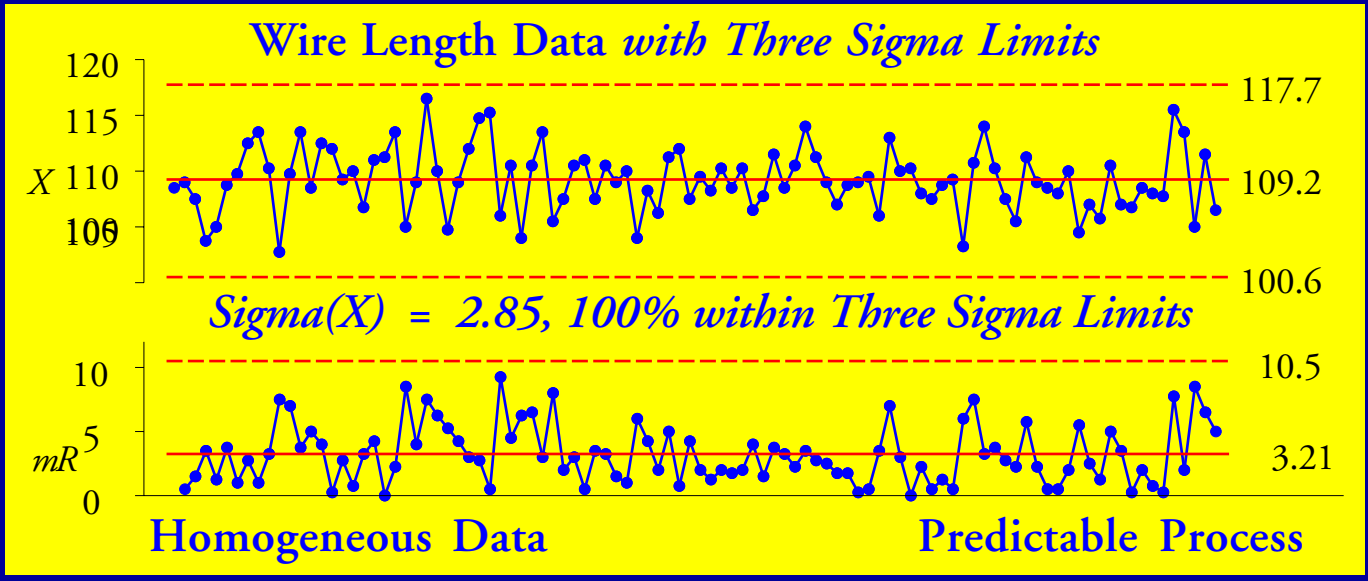


In practice we do not directly compare
the local and global measures of dispersion.

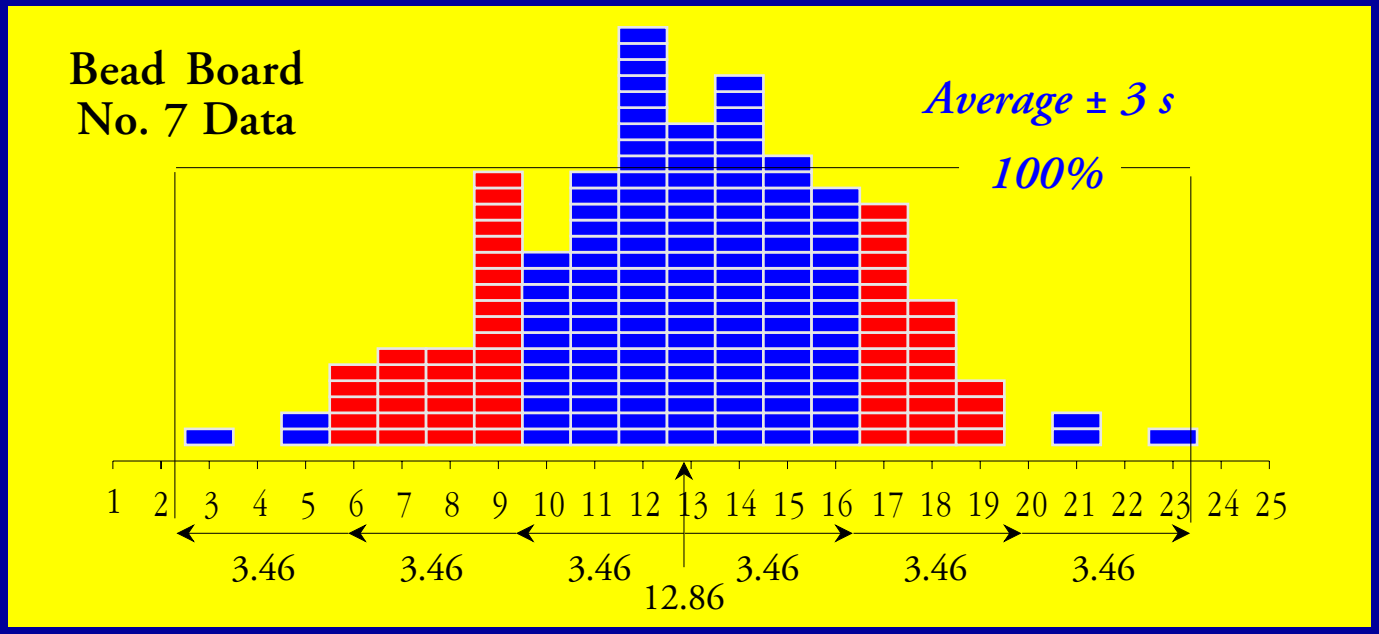
Instead we use the local measure of dispersion to compute
Three Sigma Limits
and look for points outside these limits.

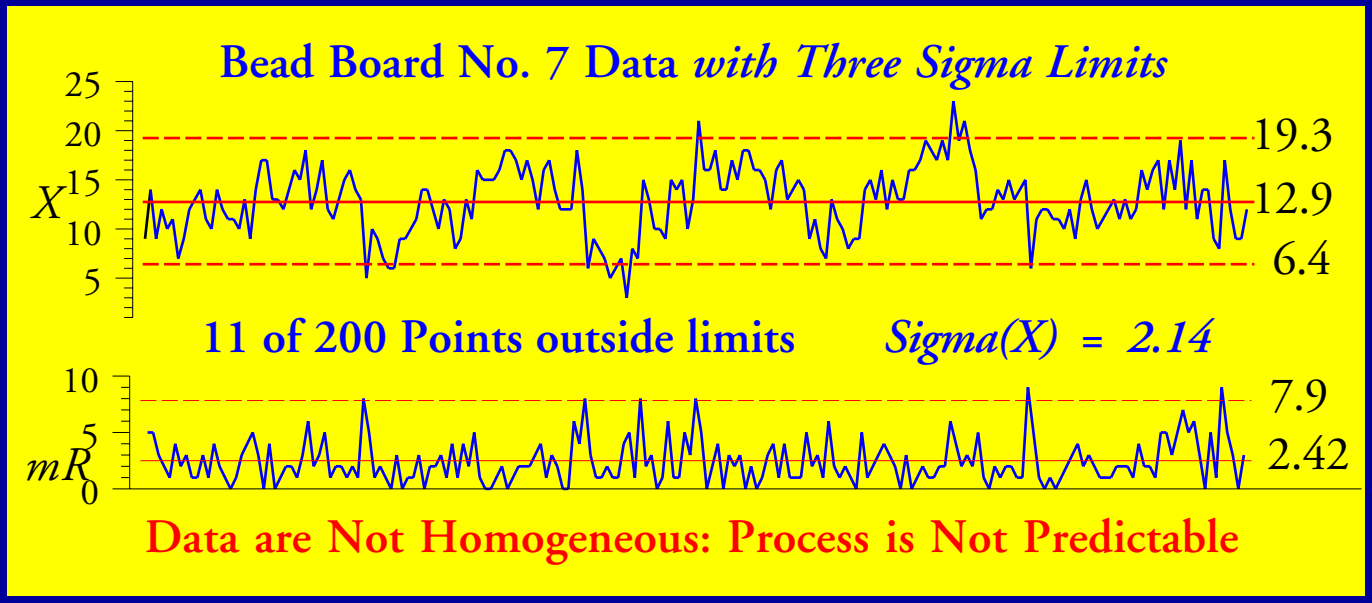
THE HOMOGENEITY QUESTION



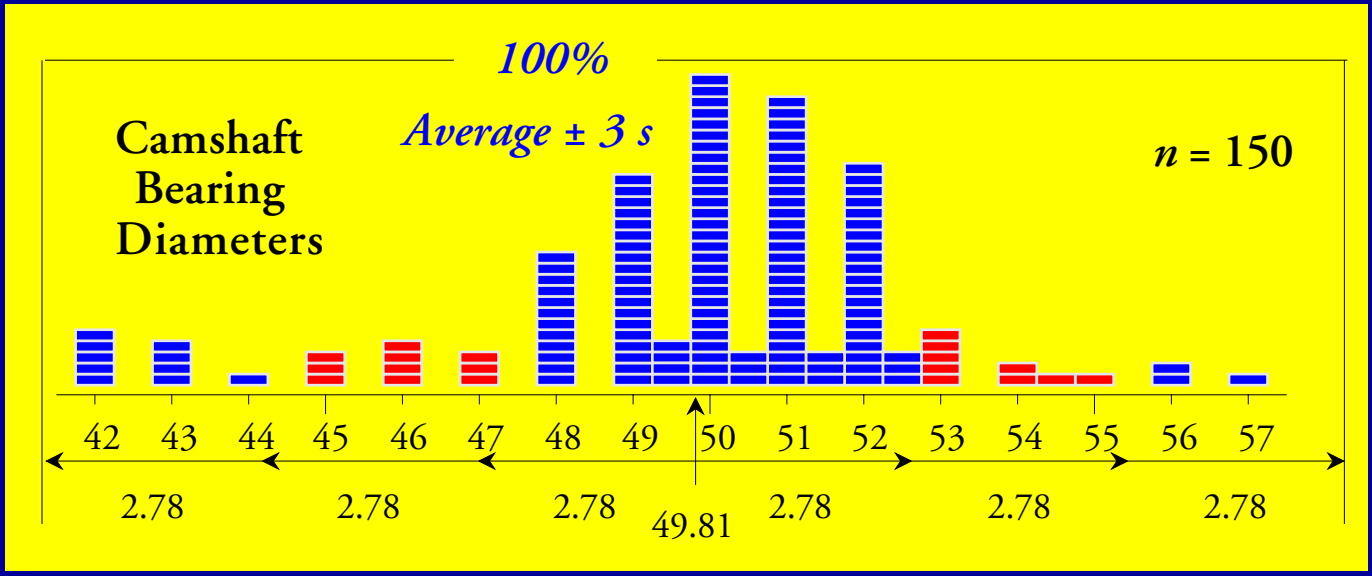


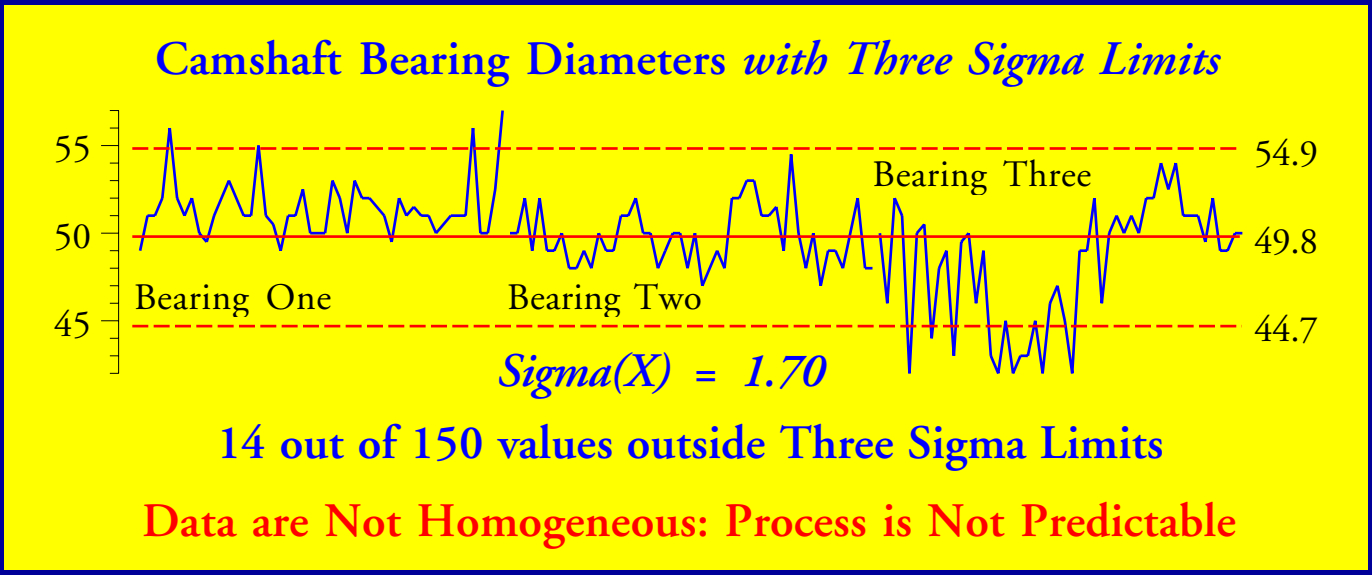
THE HOMOGENEITY QUESTION





The Homogeneity Question

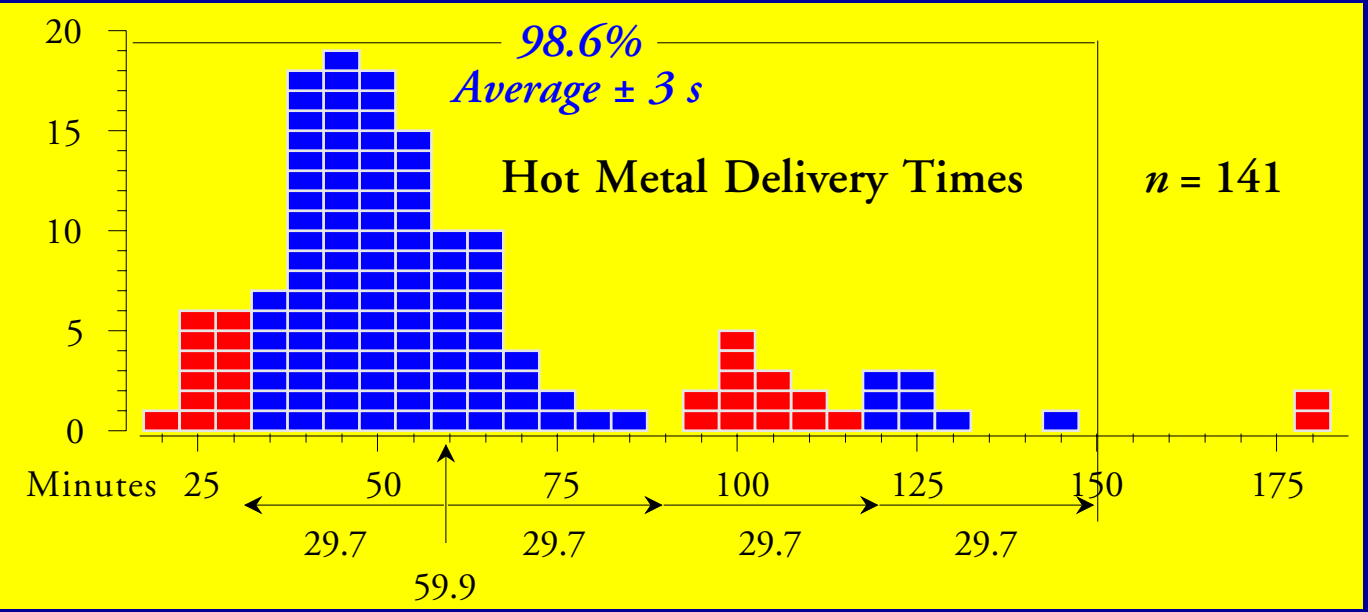


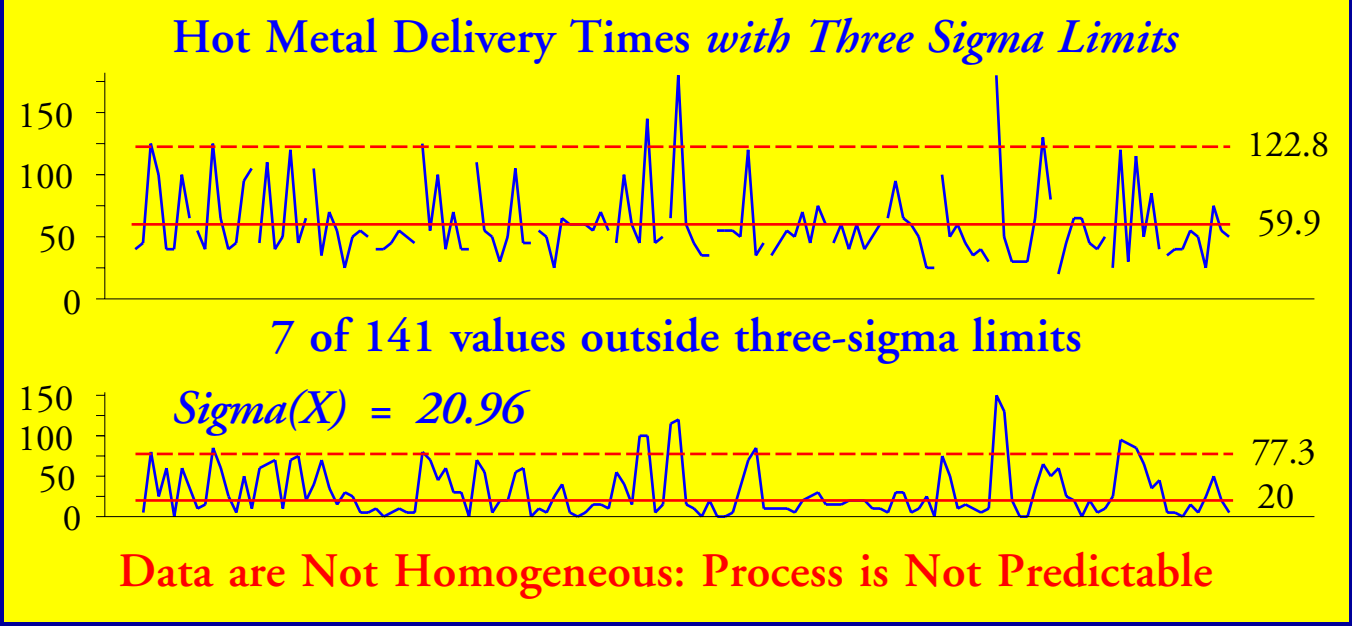


14 out of 150 values outside Three Sigma Limits

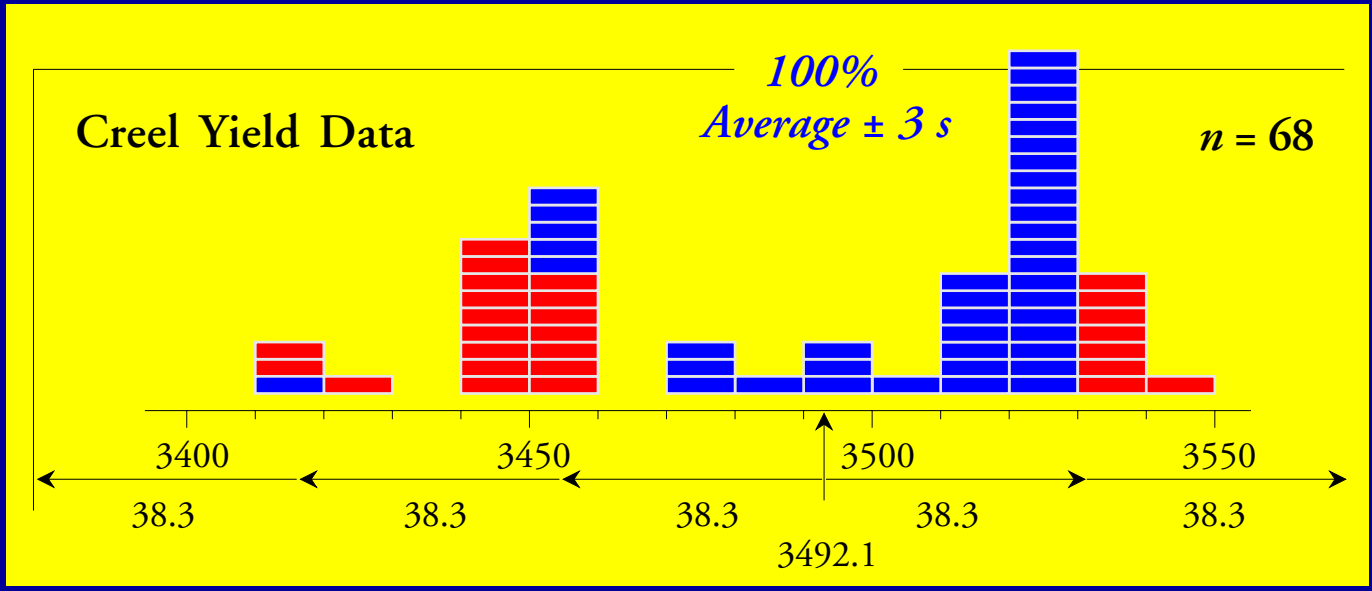
Data are Not Homogeneous: Process is Not Predictable

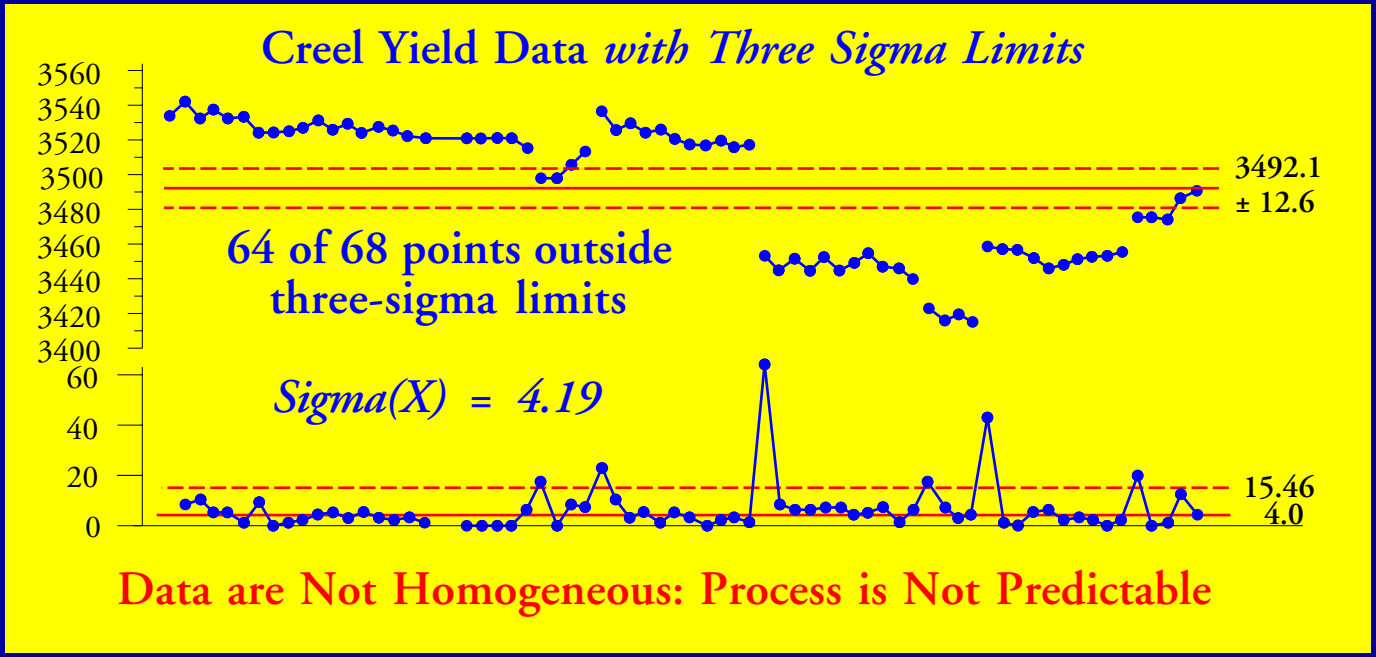
The Homogeneity Question





The Homogeneity Question





The Homogeneity Question

Only one of our six data sets was homogeneous.

Three Standard Deviation Limits bracketed *all* of the data for three of the five nonhomogenous data sets.

Three Sigma Limits detected all five of the nonhomogeneous data sets.

	Percentages Outside	
	Average ± 3 Std. Dev.	**Average ± 3 *Sigma(X)***
Wire Length Data	**0.0 %**	**0.0 %**
Bead Board No. 7 Data	**0.0 %**	**5.5 %**
Batch Weight Data	**1.1 %**	**21.6 %**
Bearing Diameters	**0.0 %**	**9.3 %**
Hot Metal Deliveries	**1.4 %**	**5.0 %**
Creel Yield Data	**0.0 %**	**94.0 %**

The only reliable way to check for homogeneity is to use a localized measure of dispersion.

The Four Questions of Statistics

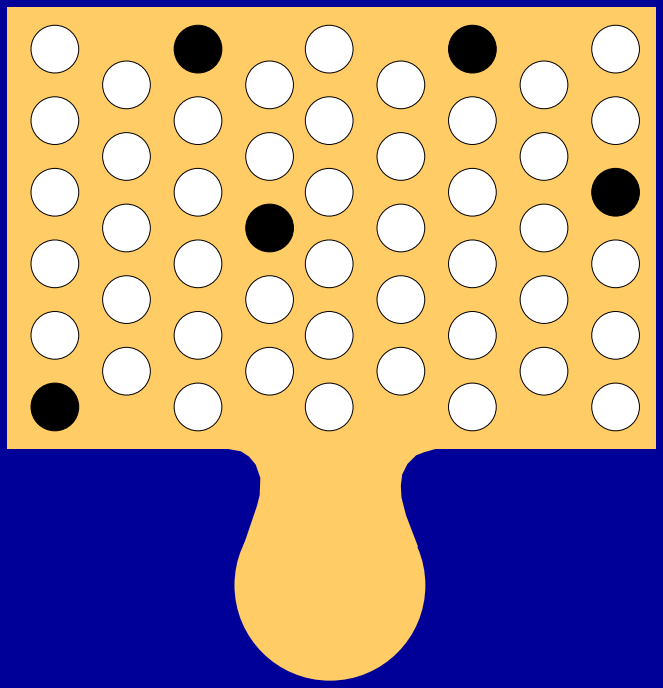
Description

Given a collection of values

is there some meaningful way

to summarize the information

using one or two summary values?

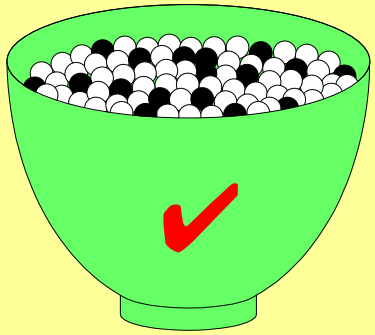


Five Black out of 50 beads
is 10 percent black.

The Four Questions of Statistics

Probability

Beginning with a Known Universe:



Urn contains
1000 Black Beads
4000 White Beads

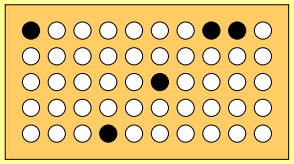
Universe
Proportion Black = 20%

Describe Various Sample Outcomes:

The chance of 1 Black in 1 Draw = 0.2

The chance of 2 Black in 2 Draws = 0.04

and so on to more complex questions:



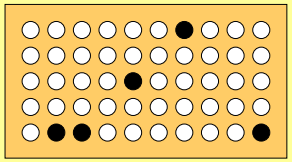
?

The probability
of getting exactly
Five Black Beads in
a Random Sample of
50 Beads is 0.030 or 3 %

The Four Questions of Statistics

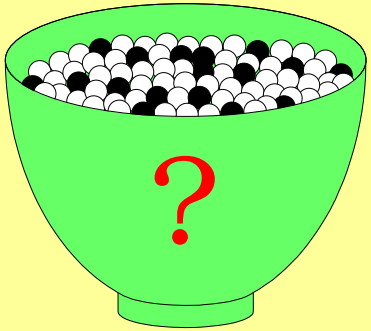
Inference

Given a Known Sample
Drawn from One Universe:



Sample contains
5 Black Beads & 45 White Beads
(10% Black)

Characterize the Unknown Universe:

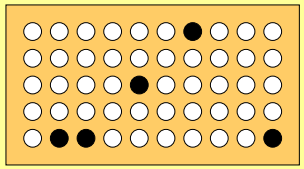


90% Interval Estimate
for
Universe Proportion Black
is
5.4% to 20.5%

The Four Questions of Statistics

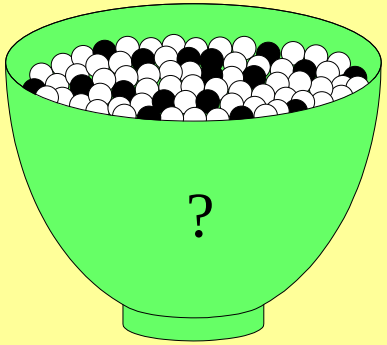
Homogeneity

Given a Collection of Data:



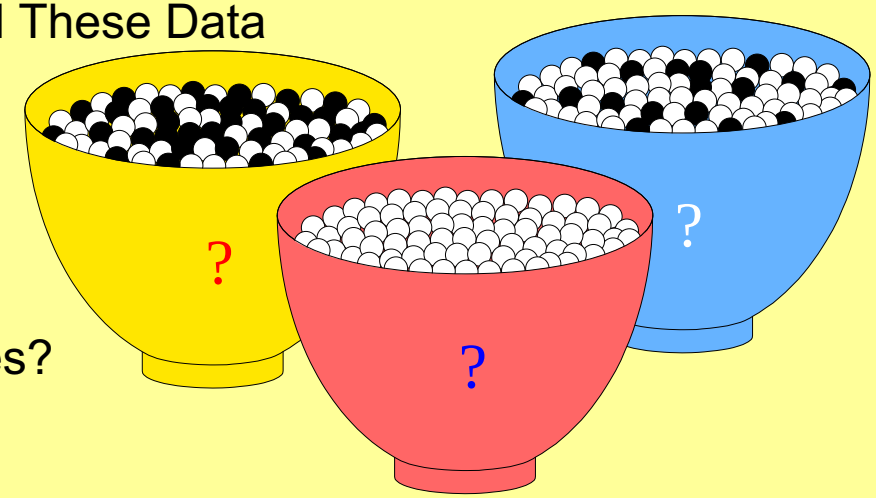
✔

Data consists of
5 Black Beads &
45 White Beads
(10% Black)

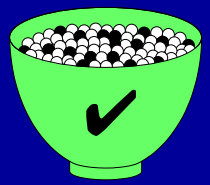


Did These Data
Come From
One Universe?

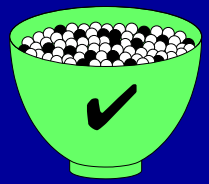
Or Could These Data
Have
Come
From
Many
Different
Universes?



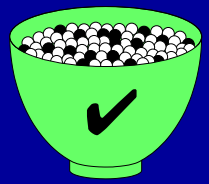
The Four Questions of Statistics



Descriptive statistics assume
the data come from a single universe.

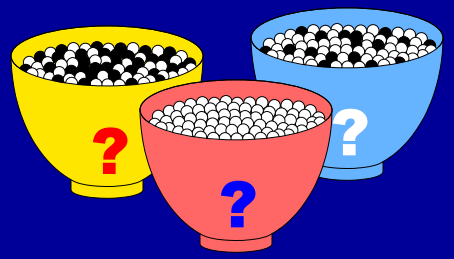


Probability theory assumes
the data come from a single universe.



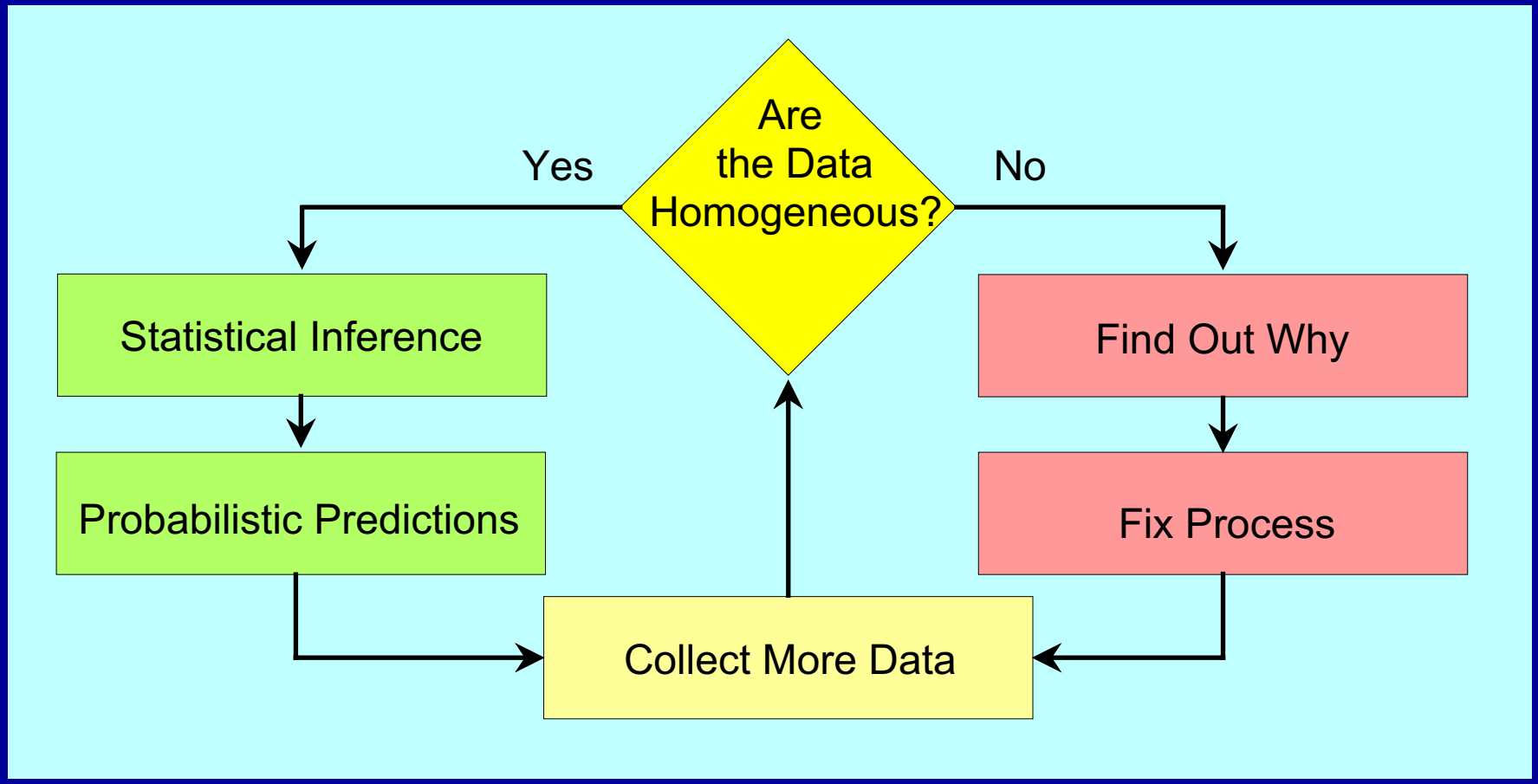
Statistical inference assumes that you have a sample
that is known to have come from a single universe.

When your data come from multiple universes
descriptive statistics fail,
probability theory fails,
and statisitcal inference
becomes meaningless.



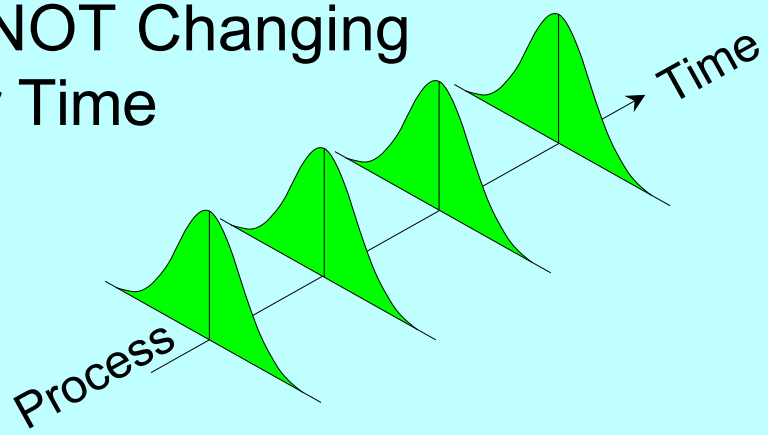
The Essence of Data Analysis

The question of homogeneity is primary



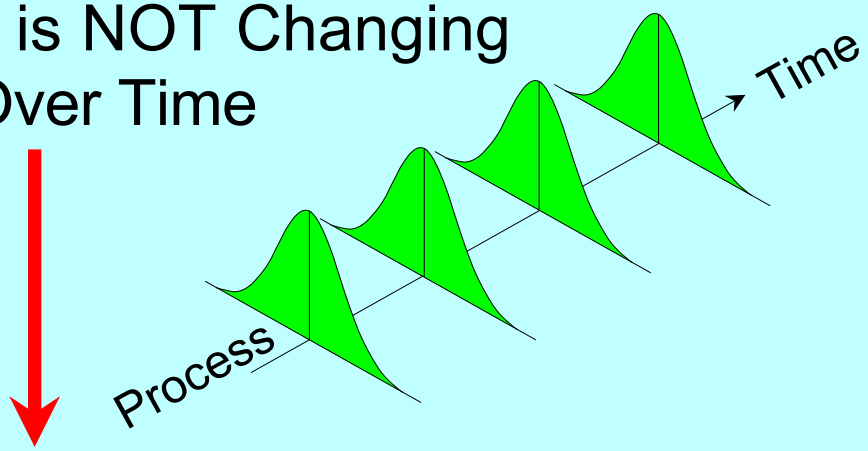
Computing Statistics & Estimating Parameters

1. When the Underlying Process is Predictable it is NOT Changing Over Time

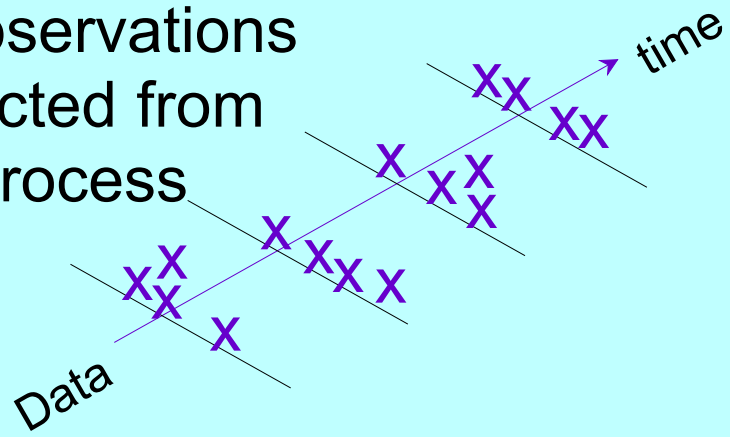


Computing Statistics & Estimating Parameters

1. When the Underlying Process is Predictable it is NOT Changing Over Time

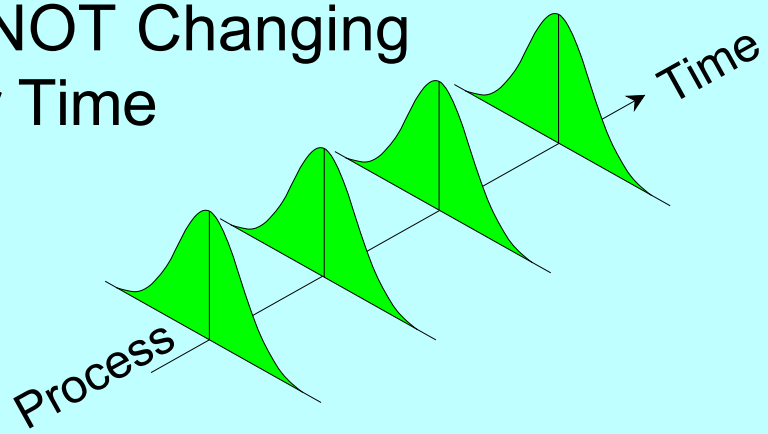


2. Observations Collected from the Process

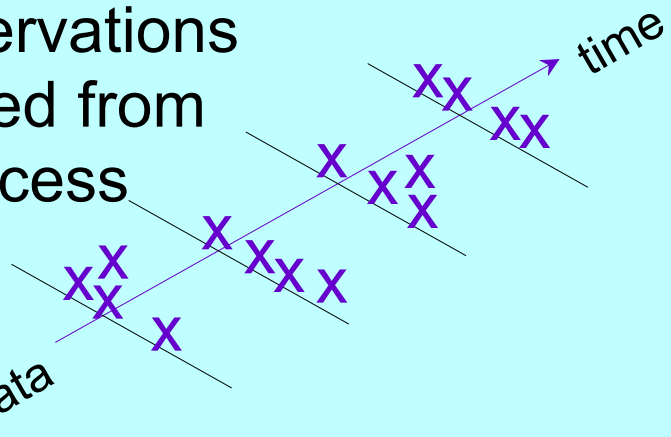


Computing Statistics & Estimating Parameters

1. When the Underlying Process is Predictable it is NOT Changing Over Time



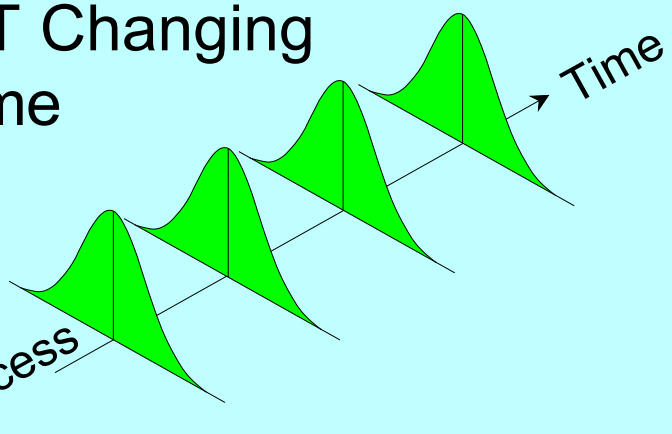
2. Observations Collected from the Process



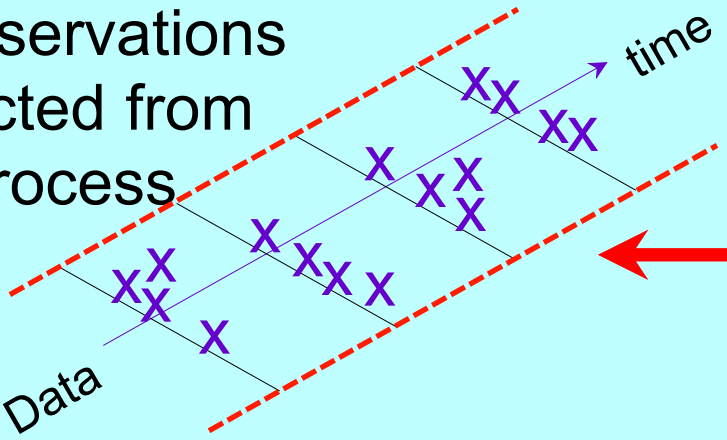
3. May be Used to Compute Summary Statistics: → Average & Average Range

Computing Statistics & Estimating Parameters

1. When the Underlying Process is Predictable it is NOT Changing Over Time



2. Observations Collected from the Process



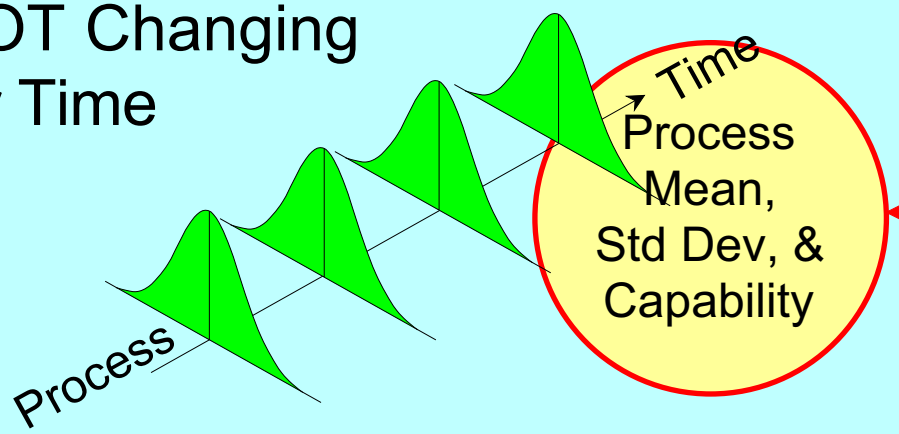
3. May be Used to Compute Summary Statistics:

Average & Average Range

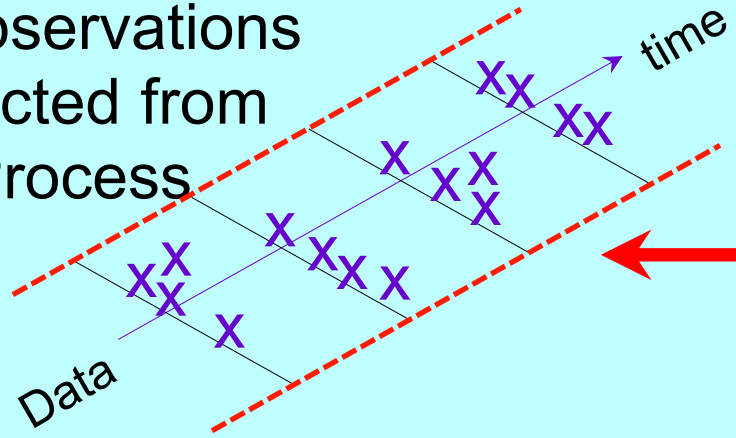
4. Which Can Be Used to Describe How Data from a Predictable Process Should Behave

Computing Statistics & Estimating Parameters

1. When the Underlying Process is NOT Changing Over Time



2. Observations Collected from the Process



3. May be Used to Compute Summary Statistics:

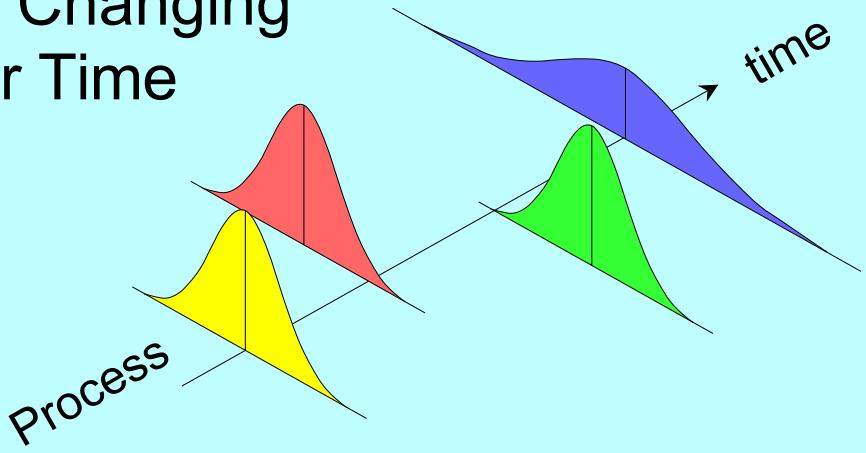
Average & Average Range

4. Which Can Be Used to Describe How Data from a Predictable Process Should Behave

5. And Can Also Be Used to Estimate the Parameters of a Predictable Process

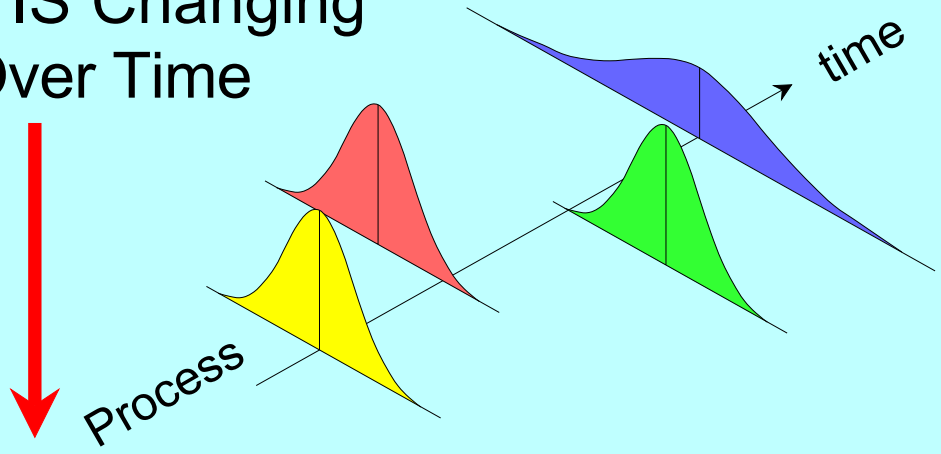
Computing Statistics & Estimating Parameters

1. When the Underlying Process is Unpredictable it IS Changing Over Time

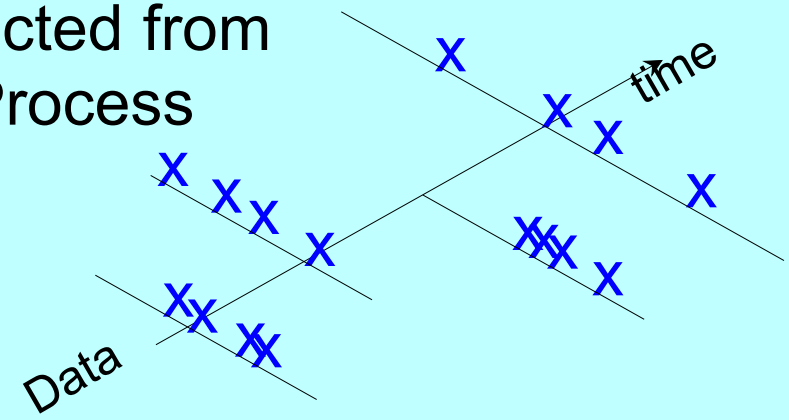


Computing Statistics & Estimating Parameters

1. When the Underlying Process is Unpredictable it IS Changing Over Time

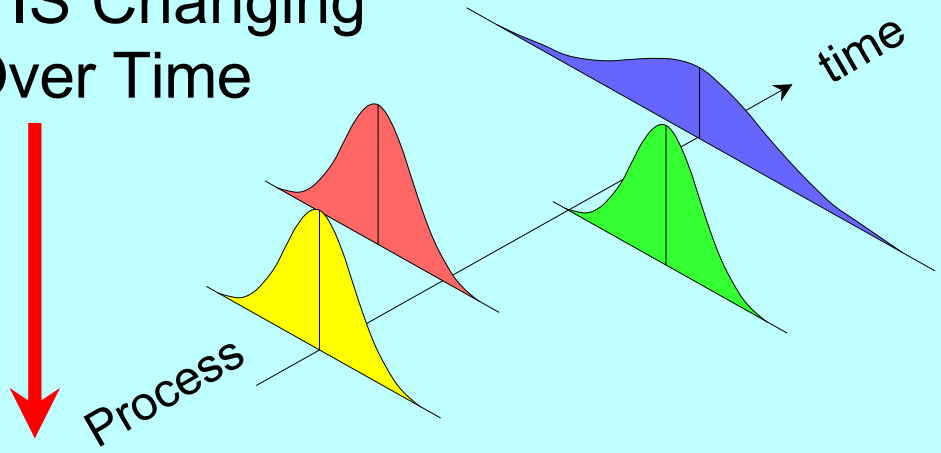


2. Observations Collected from the Process

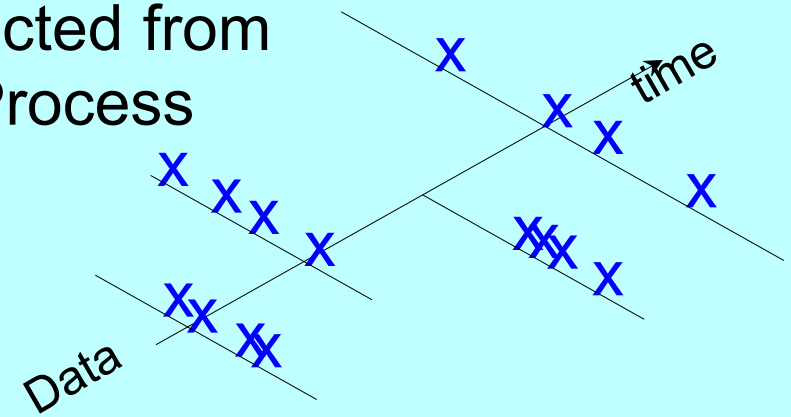


Computing Statistics & Estimating Parameters

1. When the Underlying Process is Unpredictable it IS Changing Over Time



2. Observations Collected from the Process

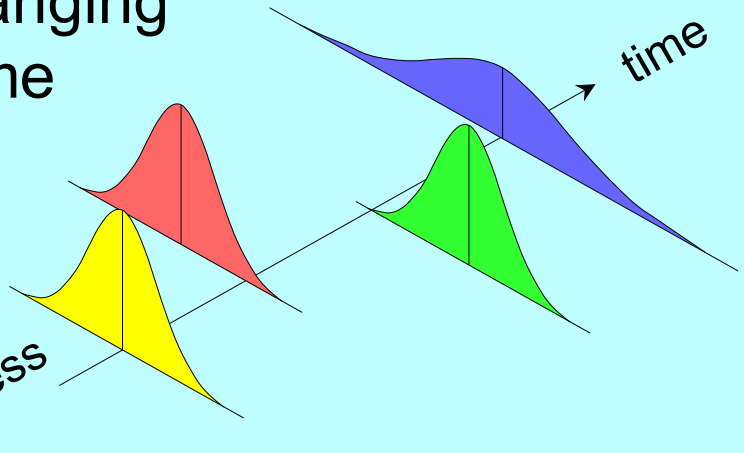


3. May still be Used to Compute Summary Statistics:

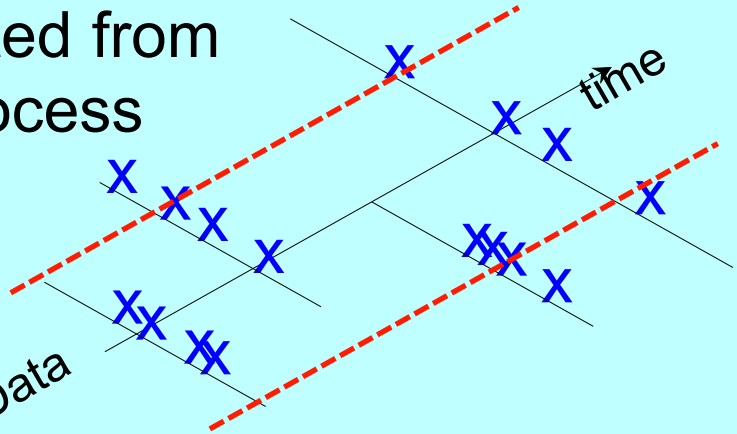
→ Average & Average Range

Computing Statistics & Estimating Parameters

1. When the Underlying Process is Unpredictable it IS Changing Over Time



2. Observations Collected from the Process



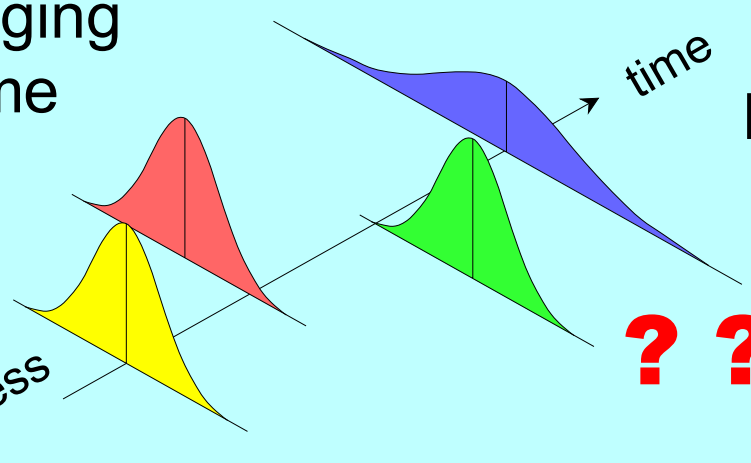
3. May still be Used to Compute Summary Statistics:

Average & Average Range

4. Which Can Be Used to Describe How Data from a Predictable Process Should Behave

Computing Statistics & Estimating Parameters

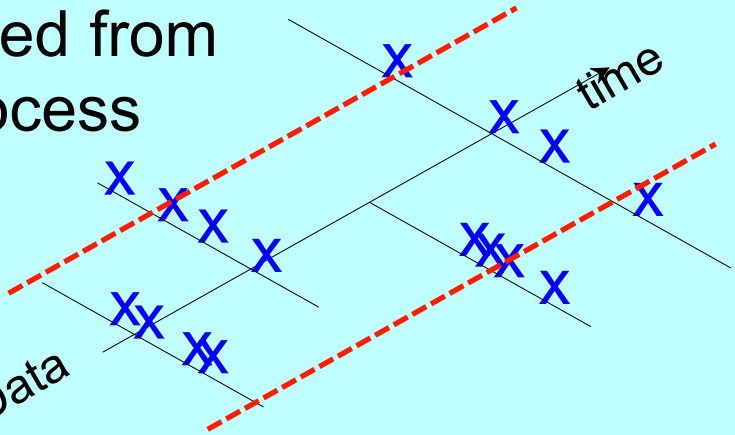
1. When the Underlying Process IS Changing Over Time



? ? ?

5. But Any Attempt to Use the Summary Statistics to Estimate Process Parameters Becomes Meaningless When the Process Mean, Std Dev, & Capability Are Constantly Changing

2. Observations Collected from the Process



4. Which Can Be Used to Describe How Data from a Predictable Process Should Behave

3. May still be Used to Compute Summary Statistics:

Average &
Average Range

How Do I Qualify My Data?

The short answer to this question is:

You do not have to qualify your data

prior to placing them on a process behavior chart!

Nevertheless, you will encounter well-meaning individuals
who will impose barriers and hurdles in the way
to placing your data on a process behavior chart.

They will do this simply because they do not understand
the difference between the traditonal statistical approach
and the approach used by Walter Shewhart
when he created the process behavior chart.

So to avoid this confusion we must discuss the difference.

The Statistical Approach to Filtering Out the Noise

With traditional statistical techniques
we begin by choosing how much noise to filter out.
(Call this value P where P is some value close to 1.00.)

Then we determine which probability model to use, *f(y)*.

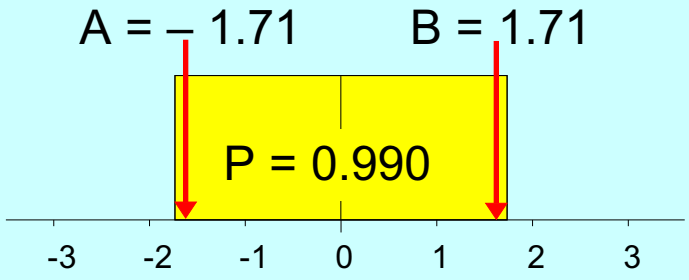
Next we find the critical values A and B
that correspond to the value P
by solving the equation for the area under a curve:

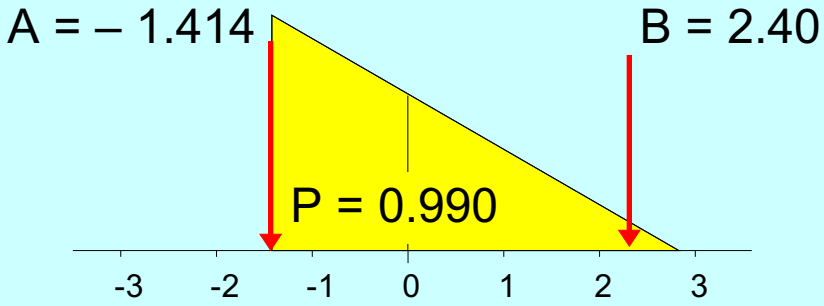
$$\int_{A}^{B} f(y)\ dy = P$$

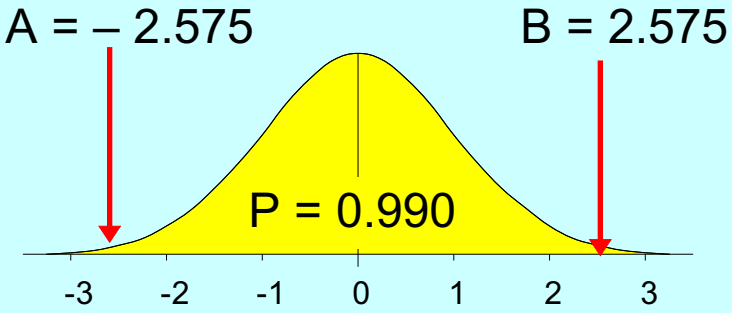
Finally, we use A and B in our analysis to filter out the noise.

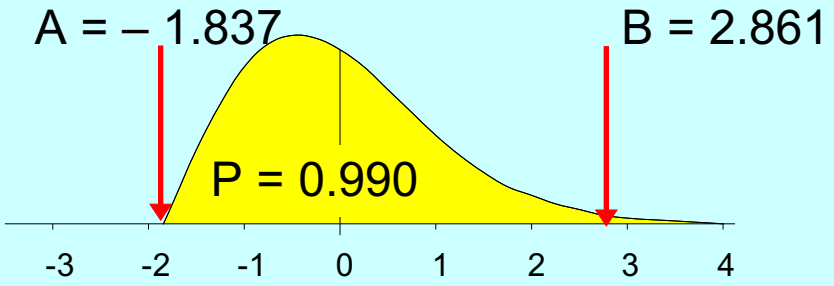
The Statistical Approach

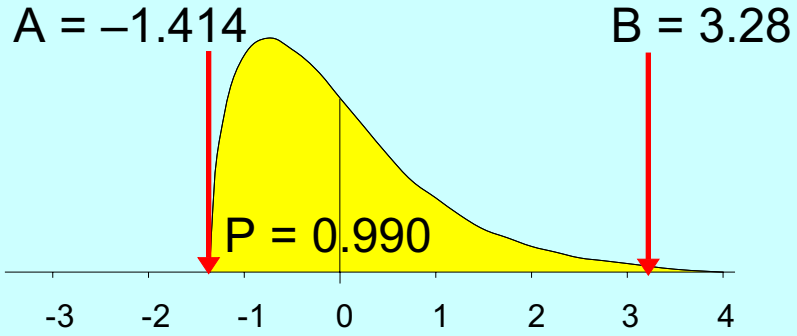
Hold P Fixed, Let A & B Vary

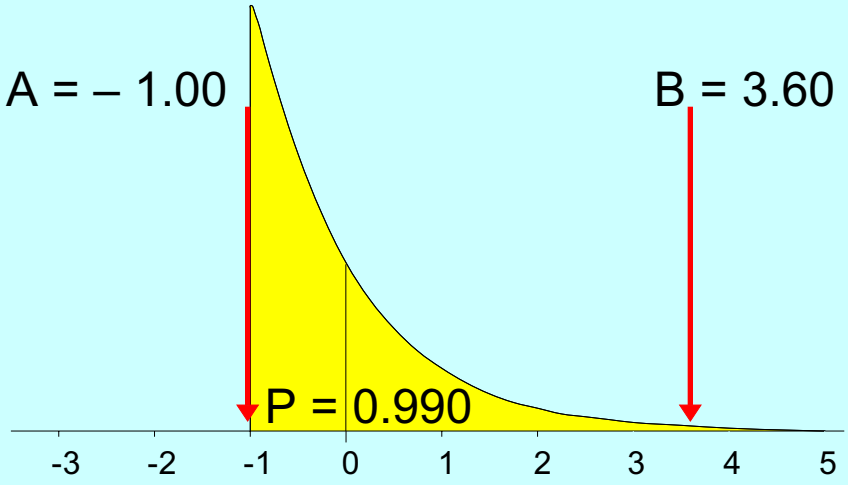












Statistics 101 in 75 Words

Hold P fixed at 0.95, 0.98, or 0.99,
then for a specific probabilty model *f(y),*
find the critical values A and B,
and use these values in your analysis.

$$\int_A^B f(y)\ dy = P$$

Statistics

Fix P near 1.0

Find A & B

The rest of your introductory class in statistics
is then concerned with how to determine
which probability model, *f(y)*, to use
for transformations of the original data such as
averages, standard deviations, proportions, etc.

Shewhart's Approach

But when working with the original data Shewhart observed:

"For the most part, however, we never know *f(y)*

in sufficient detail to set up such limits."

Shewhart

$$\int_{A}^{B} f(y)\ dy = P$$

So Shewhart approached this equation from the other end:

He chose *generic* values for A and B such that,

for *any* probability model *f(y)*,

the value for P will be reasonably close to 1.00.

Shewhart's Approach

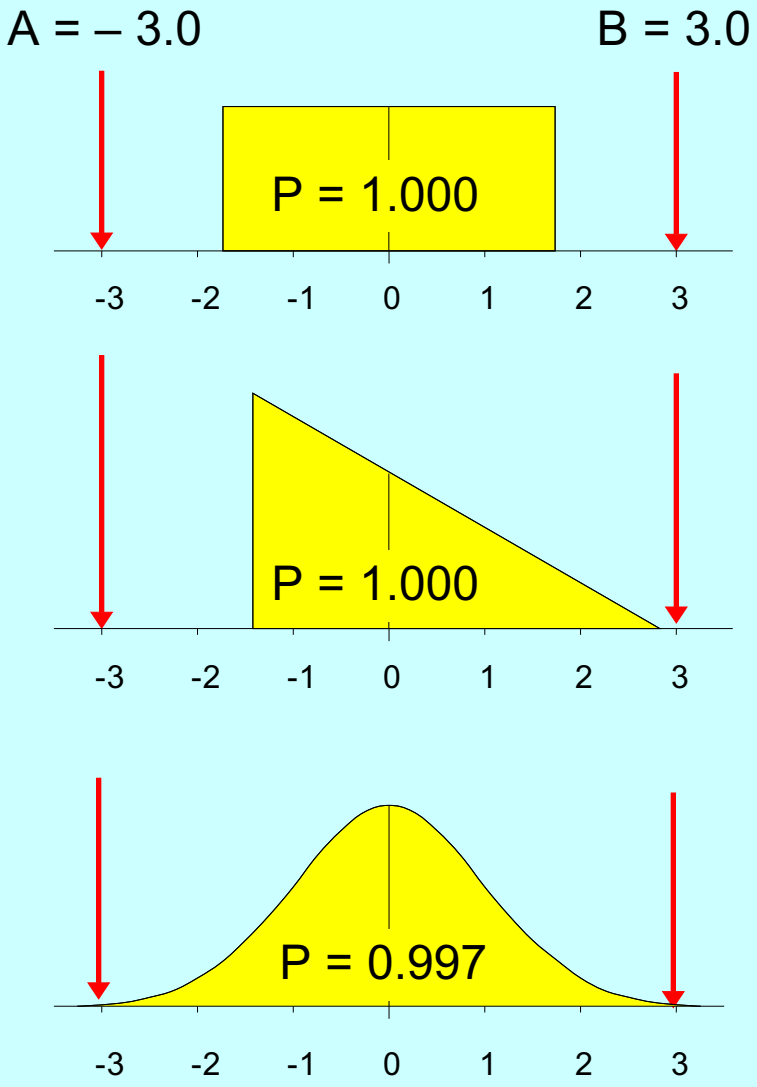
After considering the economic consequences of his approach, Shewhart chose to make the generic values A and B symmetric about the average:

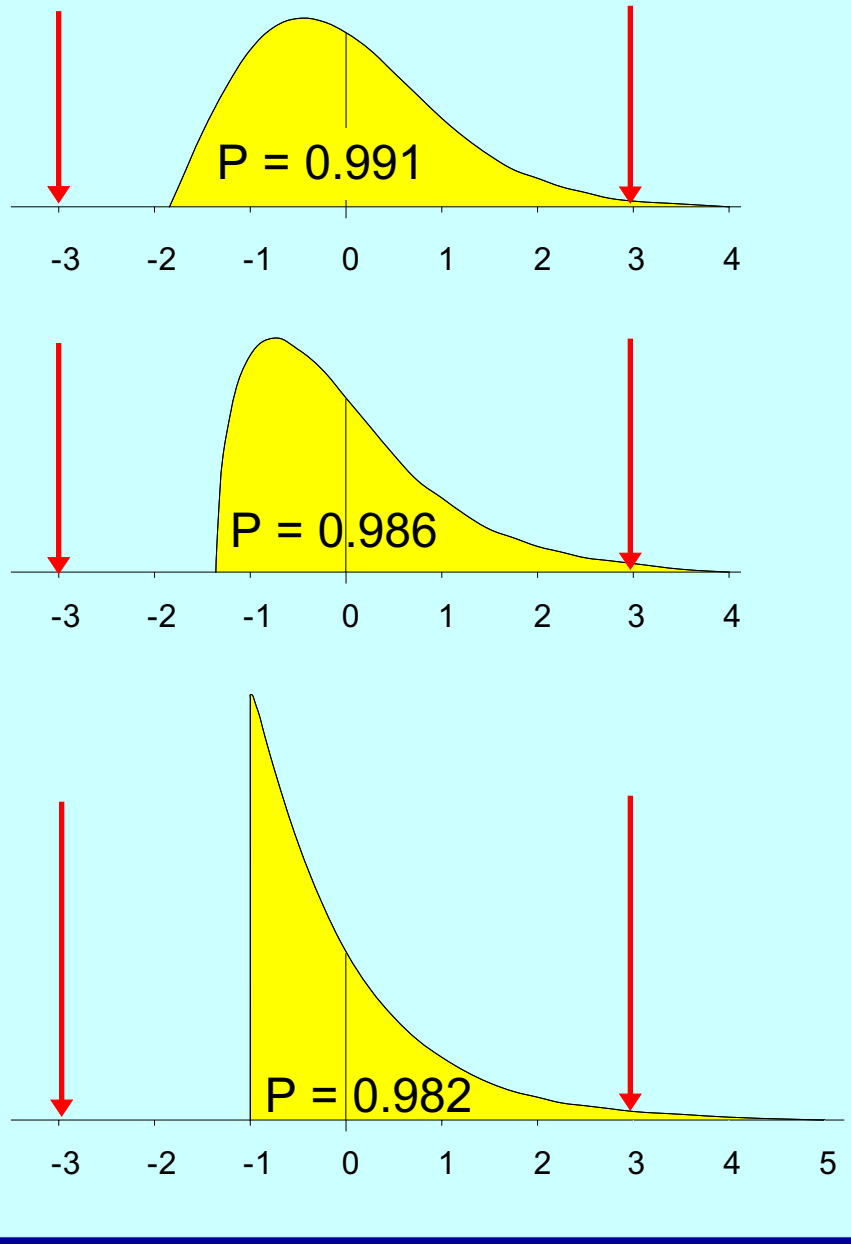
A and B = *Average* ± *t Sigma*

"We are still faced with the choice of *t*. Experience indicates that *t* = 3 seems to be an acceptable economic value."

Shewhart's Approach

Fix A&B, Let P Vary





Shewhart's Approach

As long as the generic values for A and B
result in a value for P that is reasonably close to 1.00,
the value for [1 – P] will be close to 0.00,
and false alarms will still be rare.

Thus, when an observed value for *Y*
falls outside the interval A to B,
it will still be more likely to indicate
the presence of an Assignable Cause
than to be a false alarm.

Ultimately, we do not care about the exact value of P.
As long as P is reasonably close to 1.00
we will end up making the correct decision
virtually every time.

Shewhart's Approach versus Statistical Approach

Thus, Shewhart's approach to the problem
of filtering out the noise
is the complete opposite of the statistical approach.

Shewhart

Fix A & B = ±3

$$\int_A^B f(y)\,dy = P$$

Statistics

Fix P near 1.0

Then P will be
close to 1.00
for every *f(y)*

Find unique
values for A & B
for each *f(y)*

And that is how a process behavior chart can be used
to examine the data for homogeneity
without reference to any specific probability model.

Why is this important?

Those who do not understand this difference
may seek to impose the statistical approach
because they mistake its undeniable complexity
for mathematical rigor.

But as soon as you try to hold P constant,
you will need to verify your probability model,
and all kinds of confusion follows.

Why is this important?

Out of this confusion we get novices placing obstacles

in the way of using process behavior charts:

"You must first test your data for normality."

"You might need to transform your data to make them normal."

"You cannot use auto-correlated data on a chart."

"'Phase one' charts differ from 'phase two' charts."

"Process behaviour charts will not work with skewed data."

"You need first to remove the outliers."

And many other sorts of complicated nonsense.

Why is this important?

This is why statisticians, who by nature and training
are wedded to the statistical approach,
have always been nervous with Shewhart's approach.
It just seems too simple.

Beginning with E. S. Pearson's 1935 book,
and continuing down to the present,
statisticians do not like the genreic, fixed-width limits.
They want to fine tune the limits to achieve a fixed value for P.
But as soon as they seek to do this,
they have to know which probability model to use.

Hence all of the hurdles, barriers, and prohibitions
arise out of this erroneous use of the statistical approach,
rather than being problems with Shewhart's approach.

Shewhart's Approach

Shewhart's generic, fixed-width interval
has been proven to work with all types of data.

Average ± 3 Sigma always gives P values Close to 1.00
when the process is predictable.

The simplicity of process behavior charts
does not come from overlooking the details.
It is rather the simplicity of a coherent approach
arising out of Shewhart's careful and rigorous analysis.

Beware of those who
try to make process behavior charts
more complex than they need to be.

Shewhart Quotes

"Classical statistics starts with the assumption
that a statistical universe [a probability model] exists,
whereas [SPC] starts with the assumption
that a statistical universe does *not* exist."

Transactions of the ASME, April 1943, p.224

*"We are not concerned
with the functional form of the universe,
but merely with the assumption that a universe exists."*

Statistical Method from the Viewpoint of Quality Control, p.54

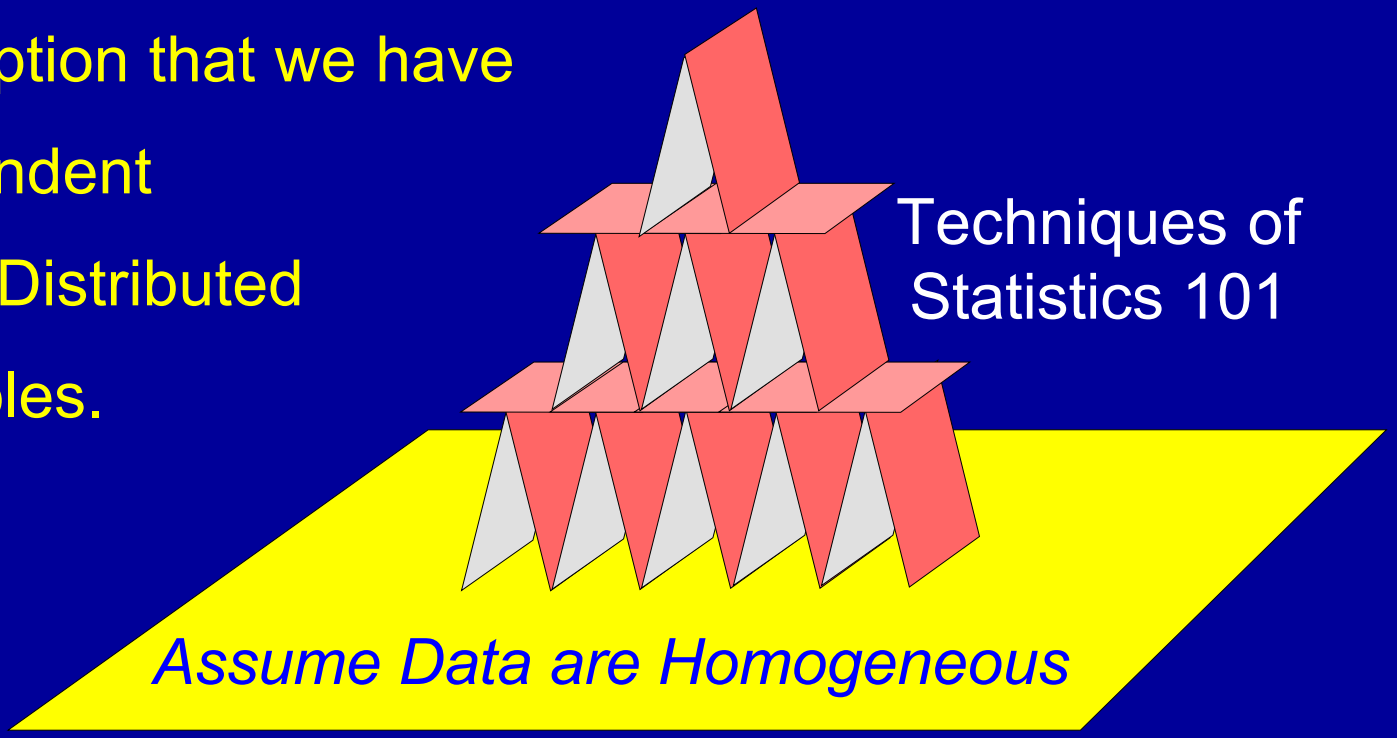
"Measureable phenomena do not obey statistical laws
until assignable causes of exceptional variation
have been found and removed from the process."

Transactions of the ASME, April 1943, p.224

www.spcpress.com/pdf/DJW350.pdf

The Secret Foundation of Statistics

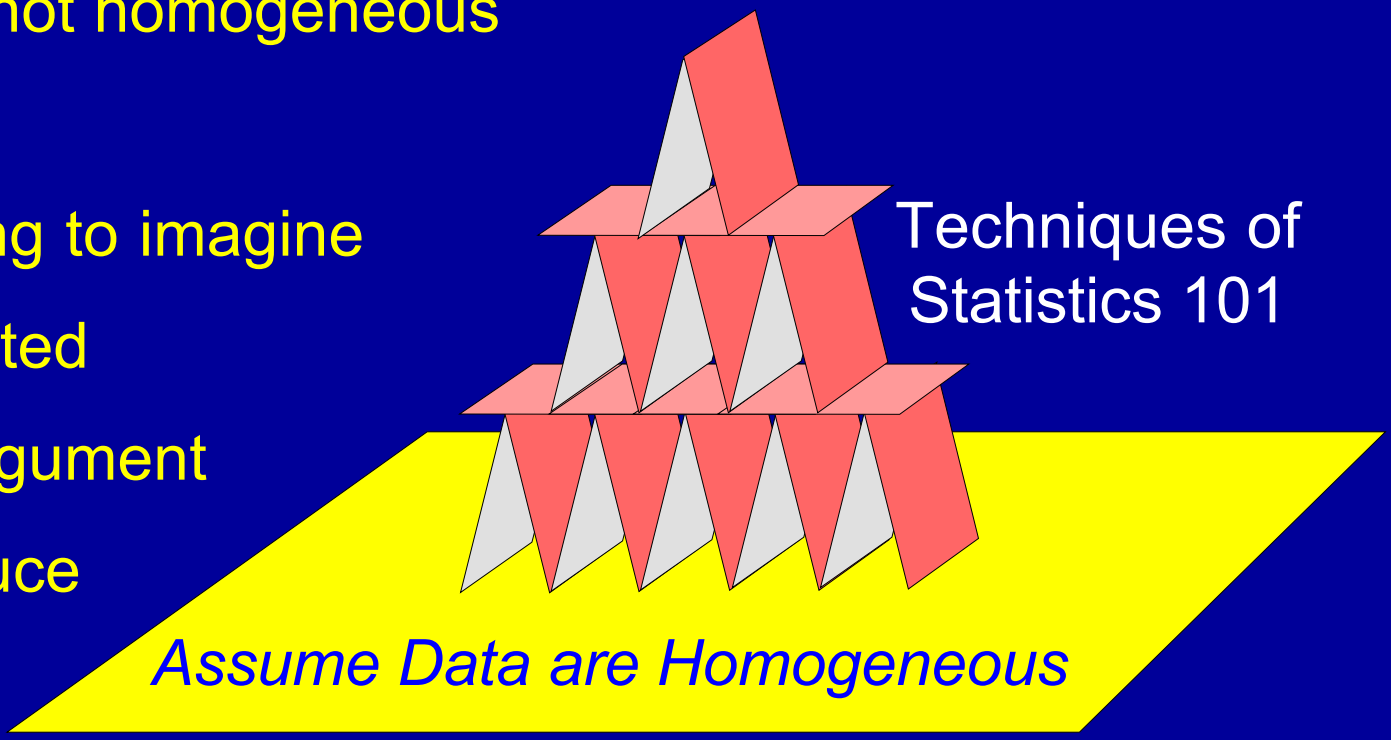
We always start our development of statistical techniques with the assumption that we have a set of Independent and Identically Distributed Random Variables.



This effectively means that we are assuming that the data will be homogeneous, and all of our techniques are built on this assumption.

The Secret Foundation of Statistics

If your data are not homogeneous
it is the height
of wishful thinking to imagine
that a sophisticated
mathematical argument
is going to produce
anything other
than nonsense.



Mere computations cannot cure a lack of homogeneity.

The Secret Foundation of Statistics

Hence, the question of homogeneity trumps all other questions.

It trumps questions about which probability model to use.

It trumps questions about

how to torture your data with transformations.

It trumps questions about which alpha level to use.

In truth, you cannot define an alpha level,

you cannot fit a probability model,

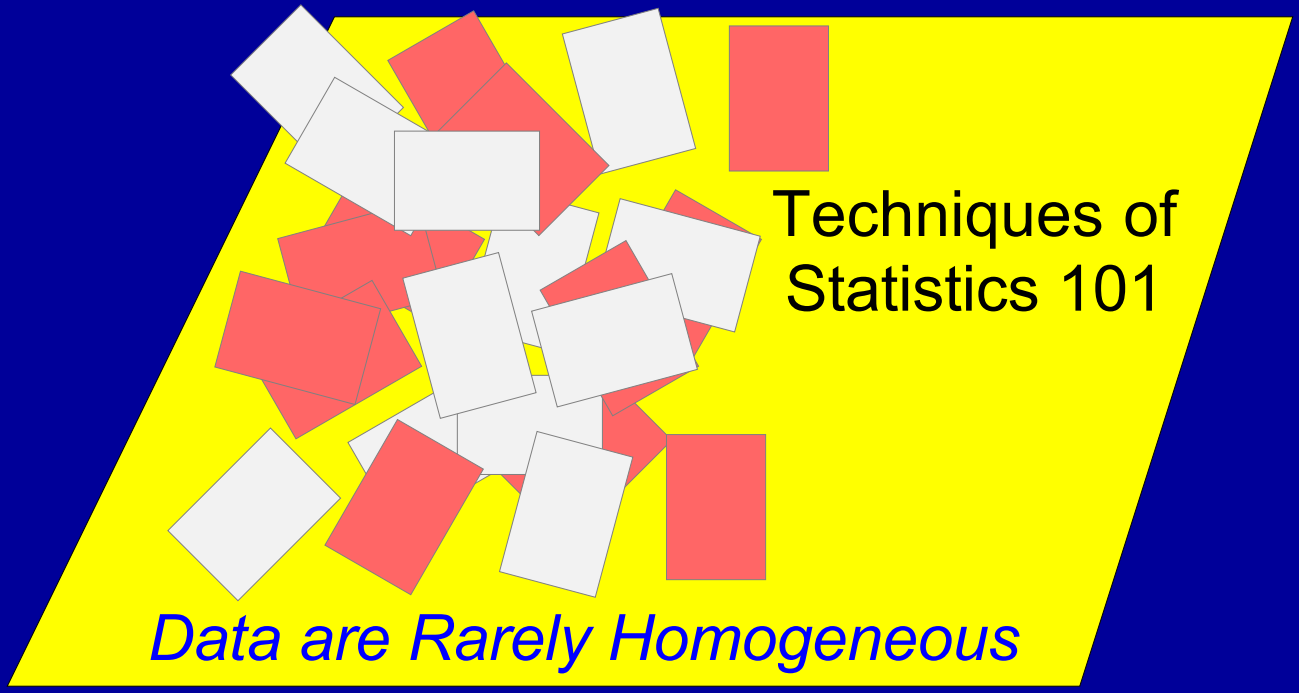
and you cannot hope

that your statistical inferences will work as advertised

if you do not have a homogeneous set of data.

The Secret of Data Analysis

The Secret of Data Analysis is that your data are rarely homogeneous.



And the easiest way to check for homogeneity before you build your house of cards is to use a process behavior chart.

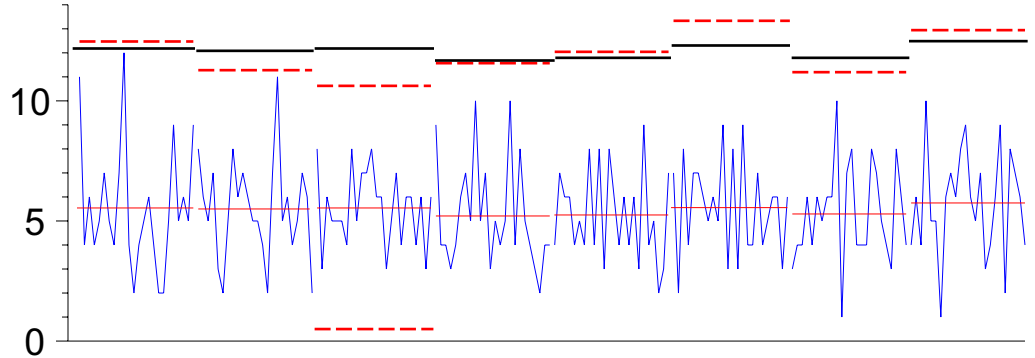


Fig. 11.3: X and np Charts for 8 Yellow Bead Experiments
8 x 24 = 192 Drawings of 50 Beads Each

X charts and np charts have same central lines
and the same running records.
Dashed Red Limits based on Average Moving Ranges.
Solid Black Limits based on Binomial Model Formulas.

Both sets of limits tell the same story for these data.

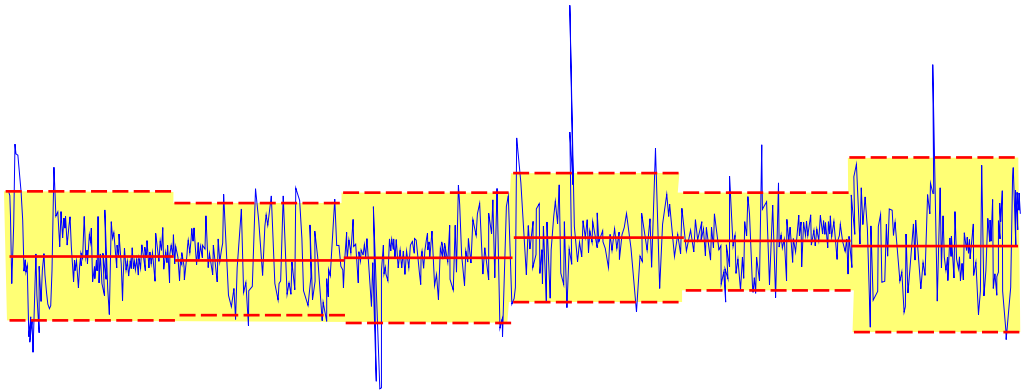
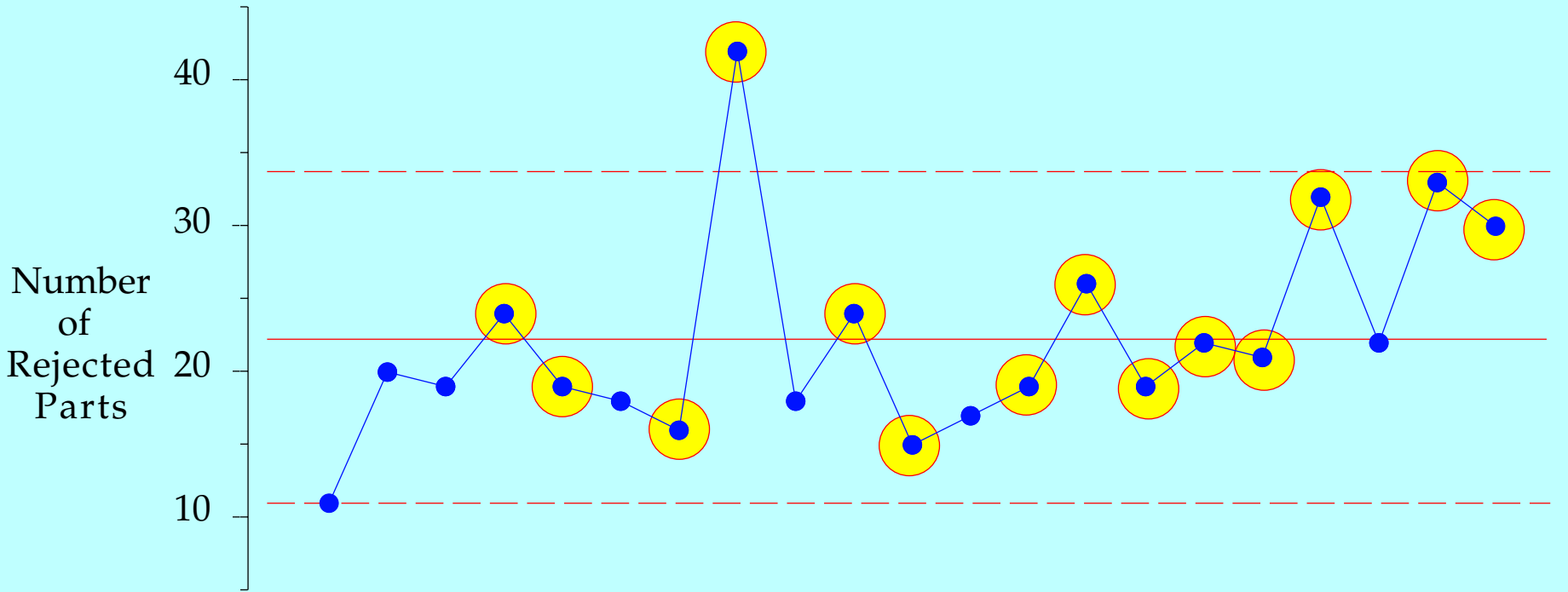


Figure 11.4: The Objective Is Not
to Compute the "Right Limitss"
But Rather to Tell the Right Story About the Process

Even though these limits vary, they all tell the same story—
this process is being operated unpredictably….
Three sigma limits are sufficient to tell the right story,
and that is why we can get good limits from bad data.
We do not have to wait until the process is well-behaved
before we compute our limits.

Sample	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*
Count	11	20	19	24	19	18	16	42	18	24	15	17	19	26	19	22	21	32	22	33	30



But the Limits Are Too Wide!

"We have to react before we get to those limits!"

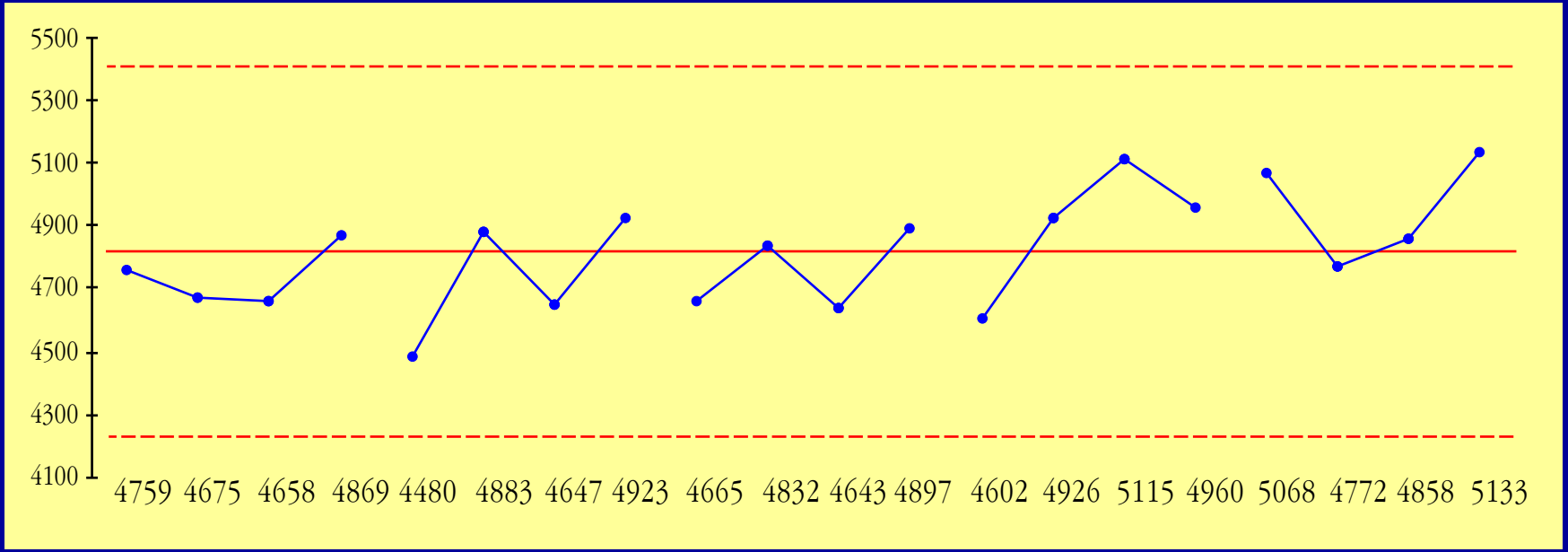
Because the XmR chart limits use the moving ranges
the limits can become inflated
(1) when the data are full of noise;
(2) when the data are full of signals; or
(3) when the data come from two or more processes.

This is why the XmR chart is intended
for use with sequences of values
that are logically comprable,
and the moving ranges should logically represent
the routine variation of the measure
plotted on the X chart.

(As described in MSD pp.141-144)

What Happens When the Data Are Full of Noise?

Total Sales for All Six Regions by Quarter for Five Years



Management Data are typically highly aggregated

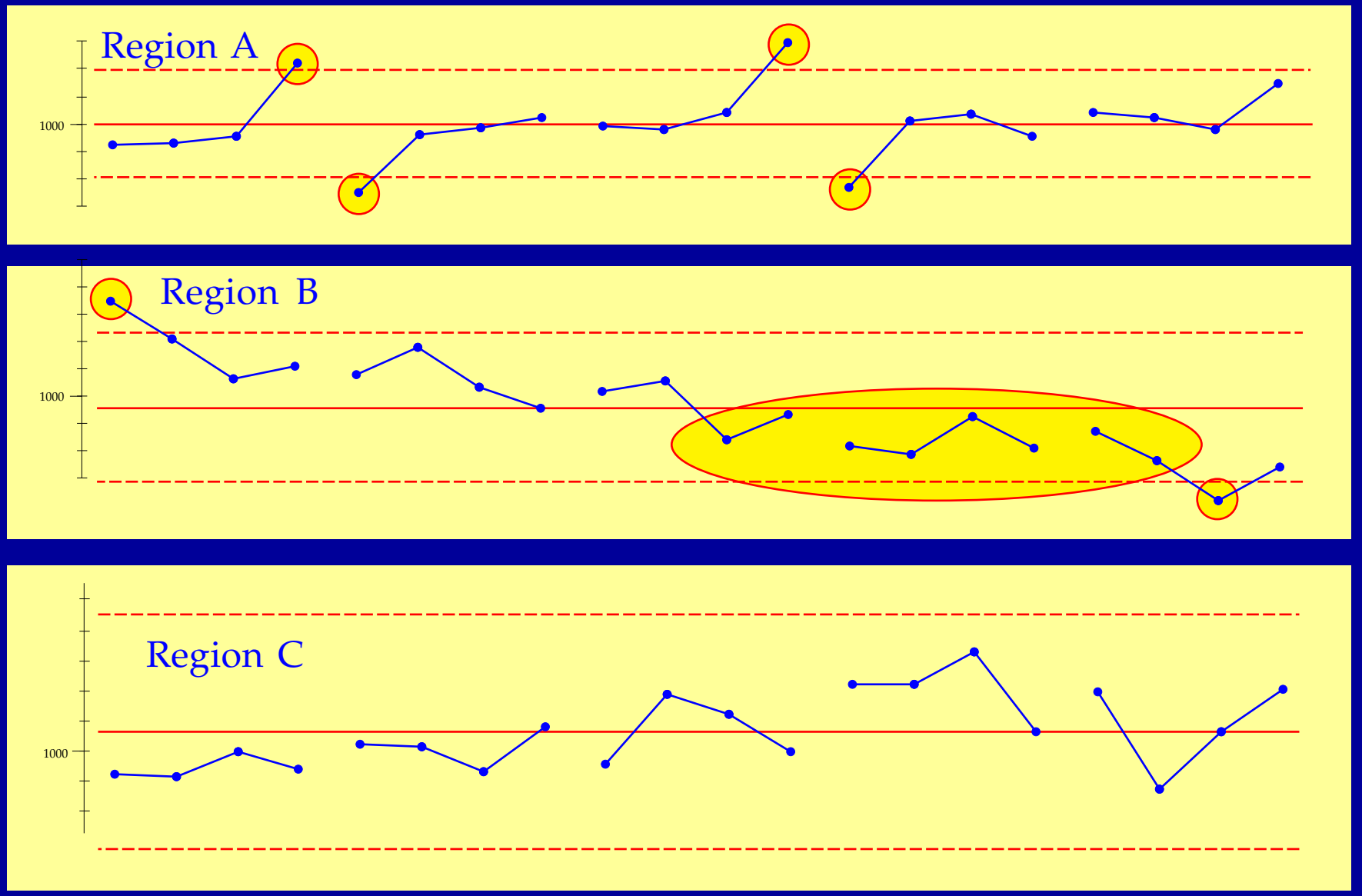
As we aggregate our data

we also aggregate the noise contained in those data.

Consider what happens when we disaggregate the data:

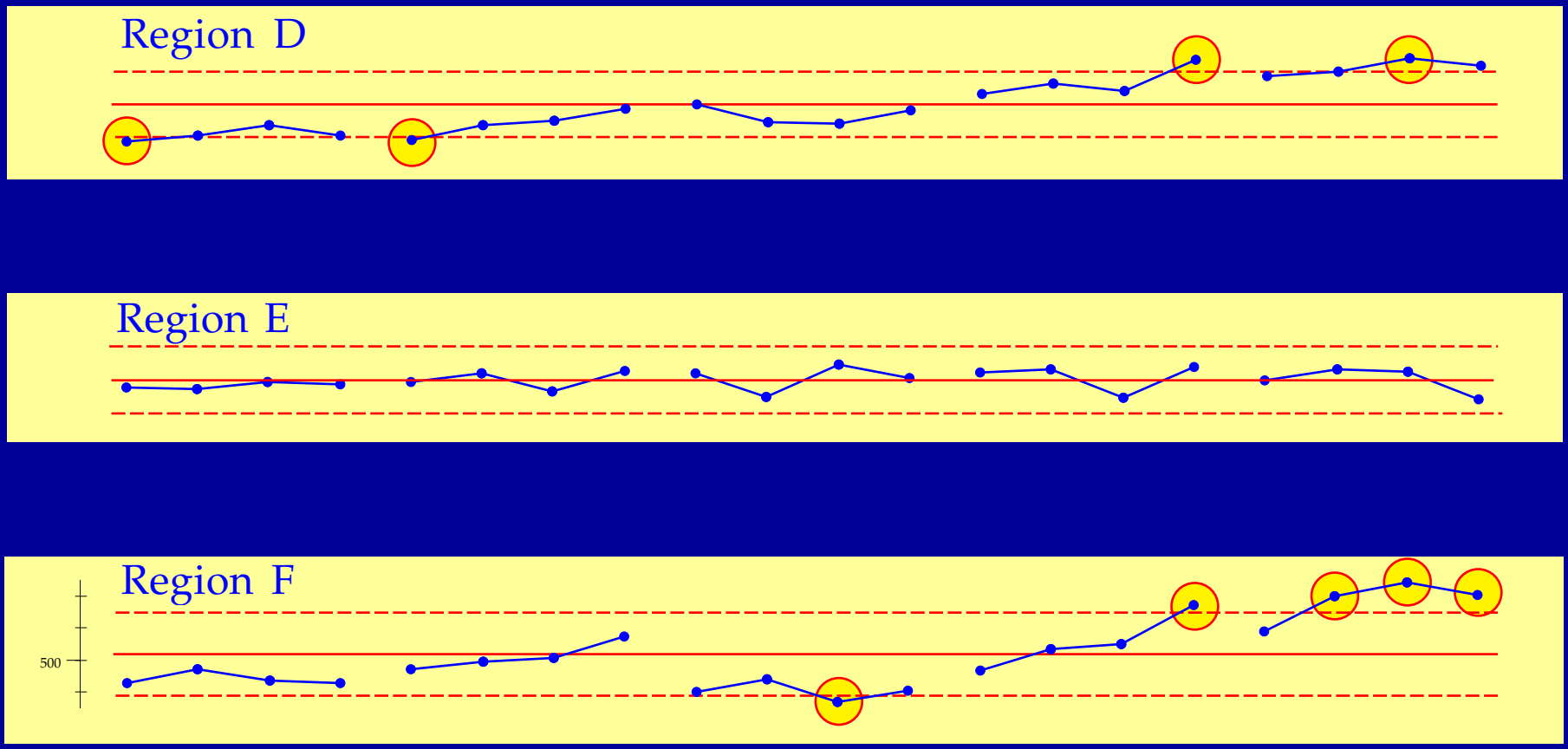
What Happens When the Data Are Full of Noise?

Quarterly Sales by Region



What Happens When the Data Are Full of Noise?

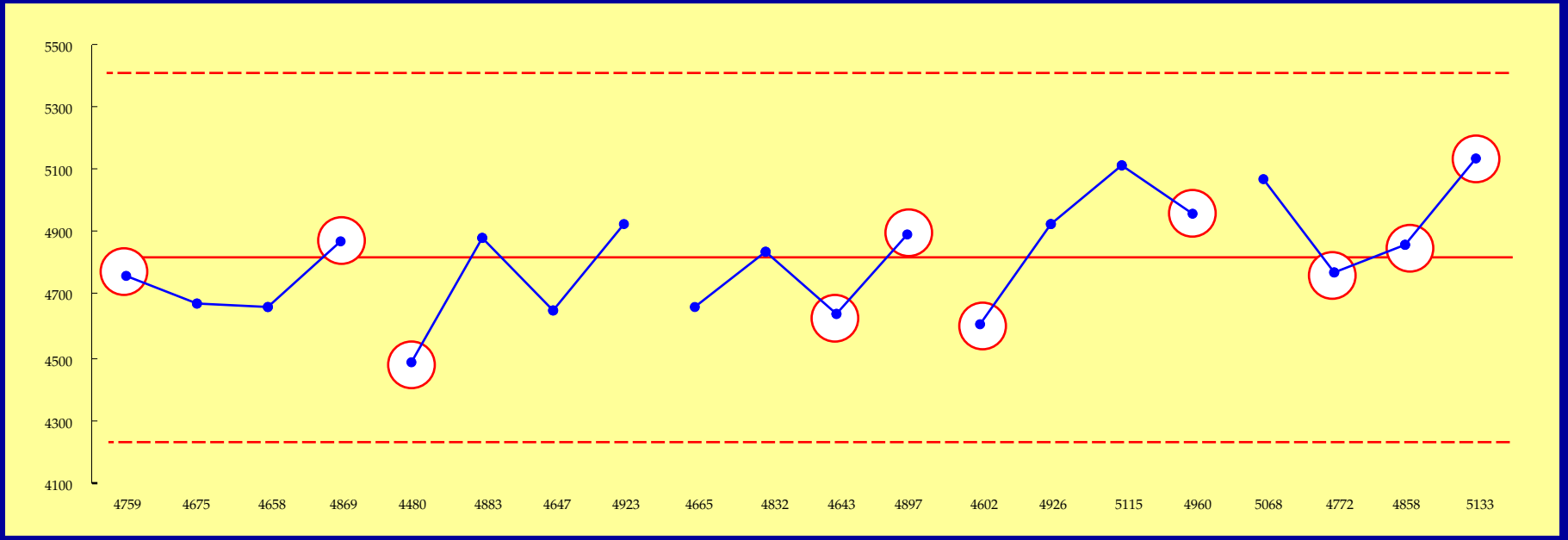
Quarterly Sales by Region



The signals will be easier to find when we disaggregate the data.

What Happens When the Data Are Full of Noise?

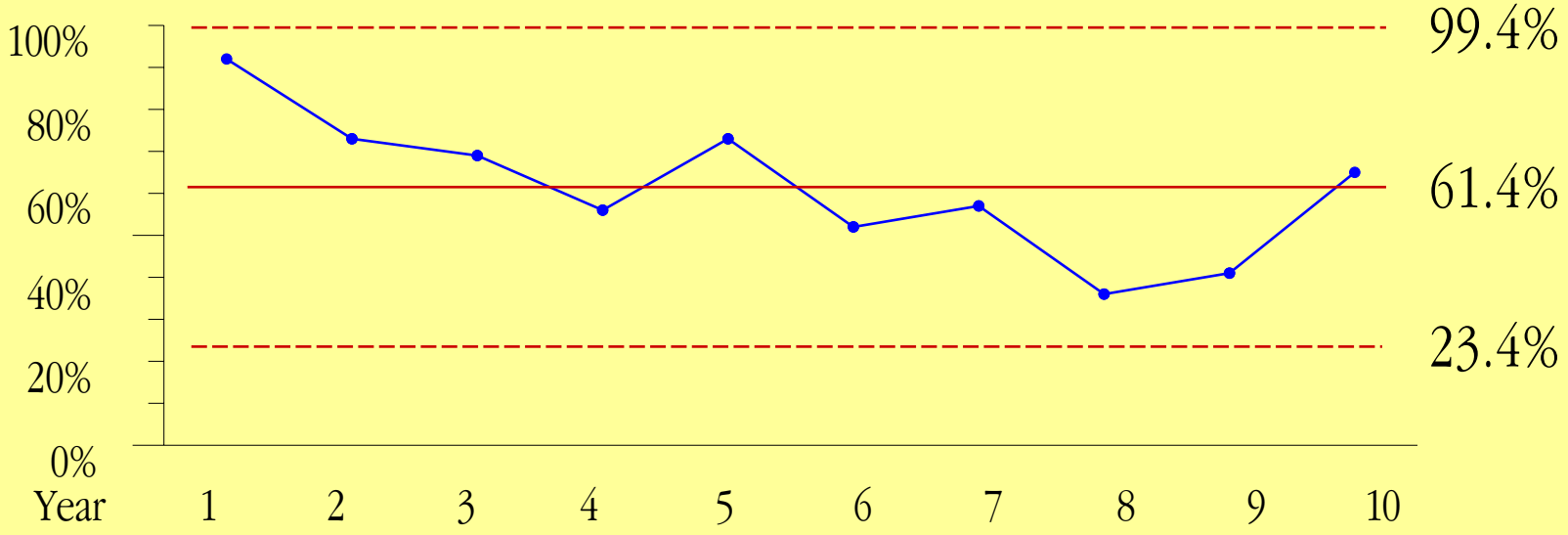
Total Sales With Rule One Signals Marked



Ten of the Twenty Quarters
were exceptional for one or more of the six regions
yet here no point even comes close
to the noise-filled limits for the aggregated totals.

What Happens When the Data Are Full of Signals?

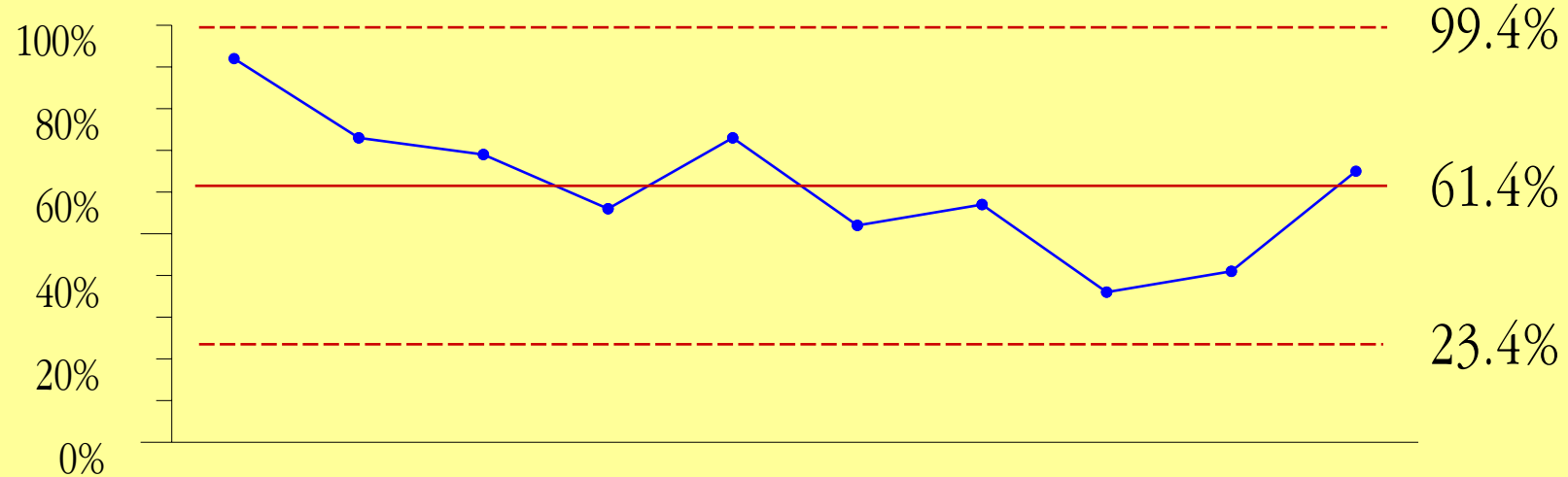
Neonatal Autopsy Percentages for Years One to Ten



The limits can become inflated
when signals become commonplace.

The assumption is that exceptional variation
will be an occasional thing
rather than a commonplace event.

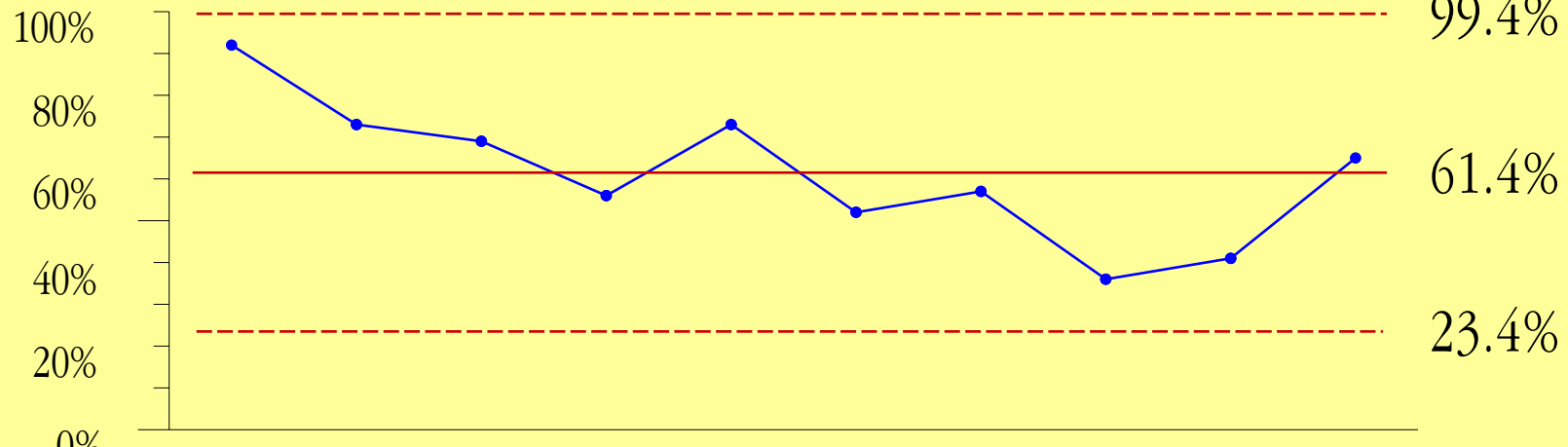
What Happens When the Data Are Full of Signals?



Year	1	2	3	4	5	6	7	8	9	10
Fatalities	38	60	67	57	49	46	53	45	32	40
Autopsies	35	44	46	32	36	24	16	16	13	26
Percent	92%	73%	69%	56%	73%	52%	57%	36%	41%	65%

No data have any meaning apart from their context.
When the annual rate goes from 92% to 36%
common sense suggests that something is changing.

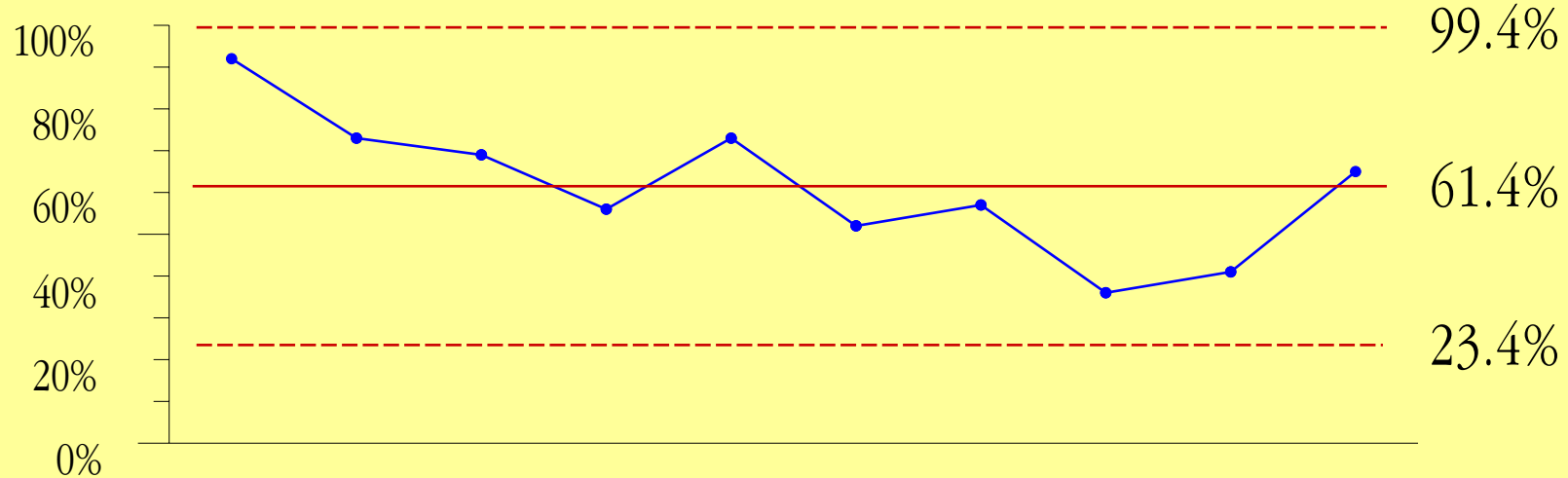
What Happens When the Data Are Full of Signals?



Year	1	2	3	4	5	6	7	8	9	10
Fatalities	38	60	67	57	49	46	53	45	32	40
Autopsies	35	44	46	32	36	24	16	16	13	26
Percent	92%	73%	69%	56%	73%	52%	57%	36%	41%	65%

The limits of 23% to 99% provide no discrimination here.
(We already knew the limits were between 0% and 100%!)
But with nothing but annual values
the year-to-year changes inflate the limits.

What Happens When the Data Are Full of Signals?

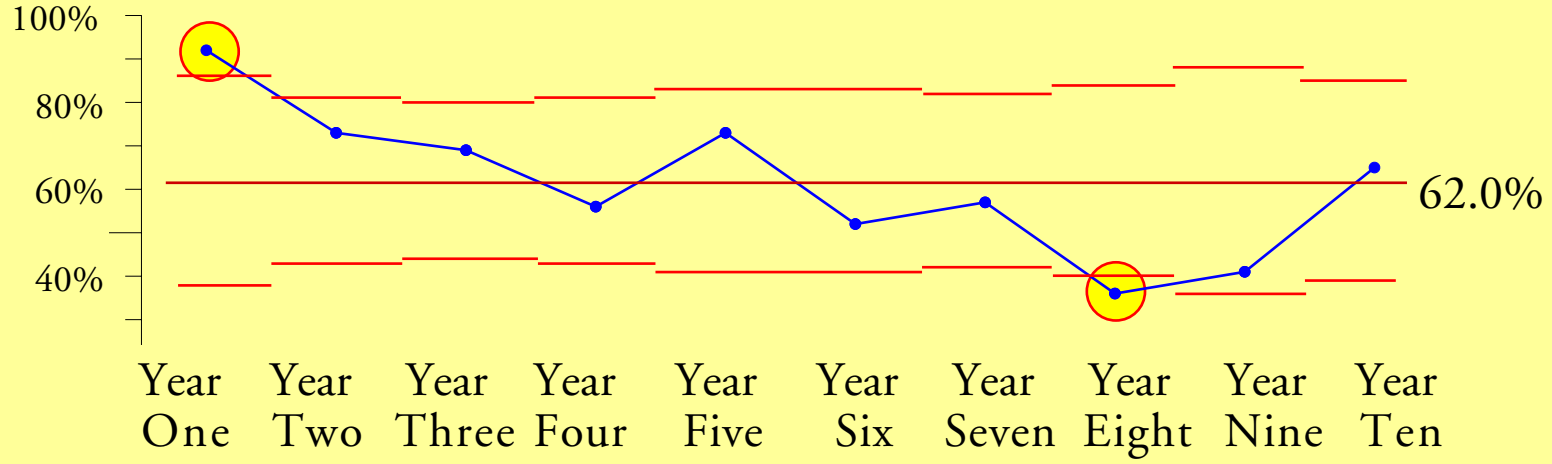


Year	1	2	3	4	5	6	7	8	9	10
Fatalities	38	60	67	57	49	46	53	45	32	40
Autopsies	35	44	46	32	36	24	16	16	13	26
Percent	92%	73%	69%	56%	73%	52%	57%	36%	41%	65%

Several authors have suggested that the wide limits here are a problem with the XmR chart approach. But it is actually a failure to make sure that successive values are logically comprable.

What Happens When the Data Are Full of Signals?

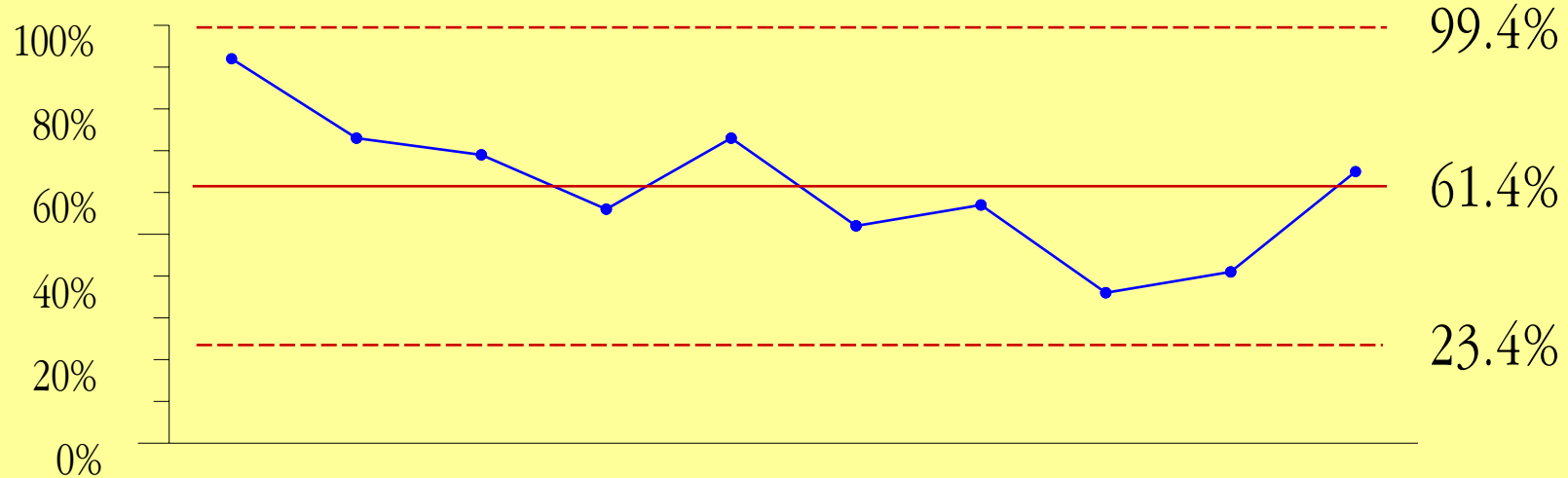
One author suggested using a p-chart for these data



While this chart does have a couple of points outside the limits, these data do not satisfy the assumptions for a p-chart.

The Binomial model requires uniformity within each time period. This means that every neonatal fatality in a given year is assumed to have the same likelihood of being autopsied. If this is not true, the counts will not follow the binomial model.

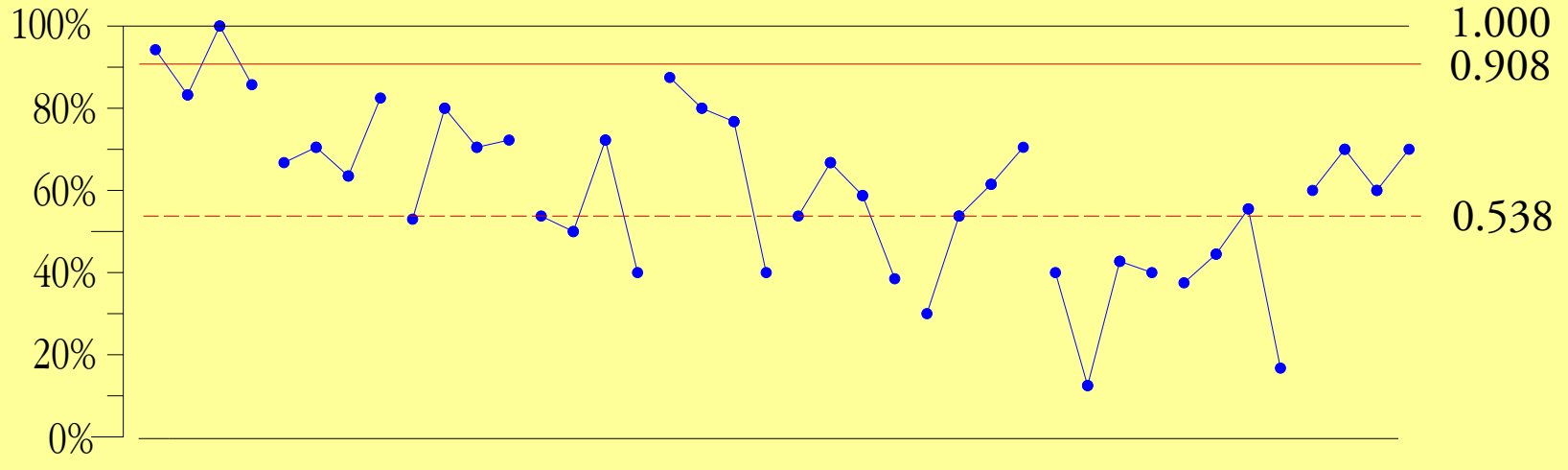
What Happens When the Data Are Full of Signals?



Year	1	2	3	4	5	6	7	8	9	10
Fatalities	38	60	67	57	49	46	53	45	32	40
Autopsies	35	44	46	32	36	24	16	16	13	26
Percent	92%	73%	69%	56%	73%	52%	57%	36%	41%	65%

Instead of going off in search of a different chart,
we need to use shorter time intervals
so that successive values will have a greater chance
of being logically similar.

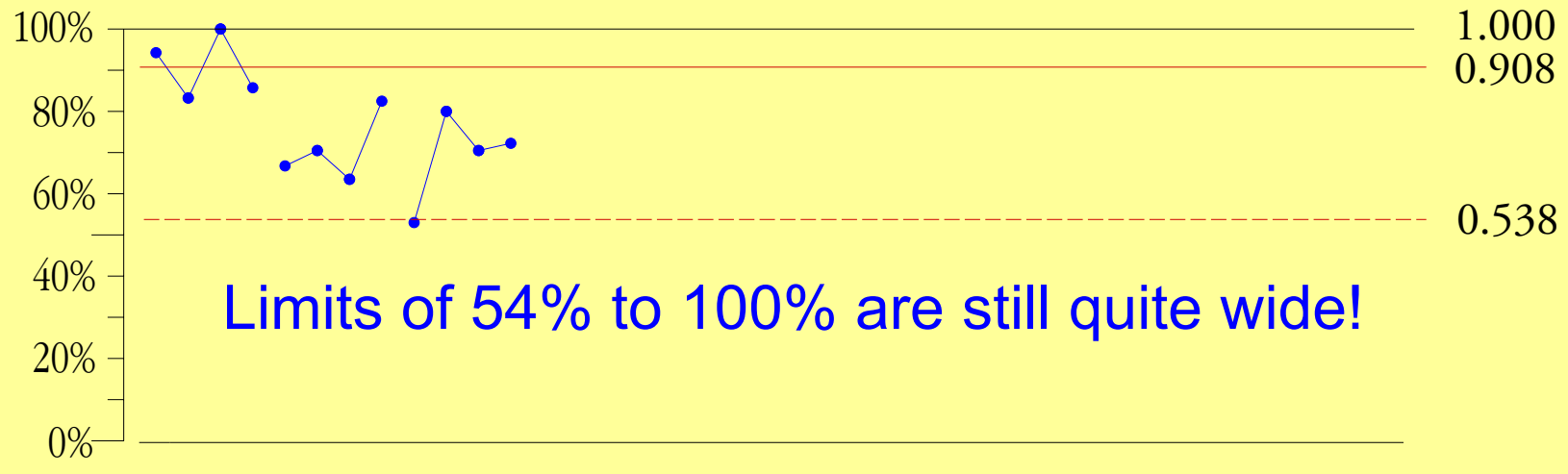
What Happens When the Data Are Full of Signals?



Year	1	2	3	4	5	6	7	8	9	10
Fatalities	38	60	67	57	49	46	53	45	32	40
Autopsies	35	44	46	32	36	24	16	16	13	26
Percent	92%	73%	69%	56%	73%	52%	57%	36%	41%	65%

Here the counts are too small
to break these data down by months,
so we break them down by quarters,
and use the first year to compute limits.

What Happens When the Data Are Full of Signals?



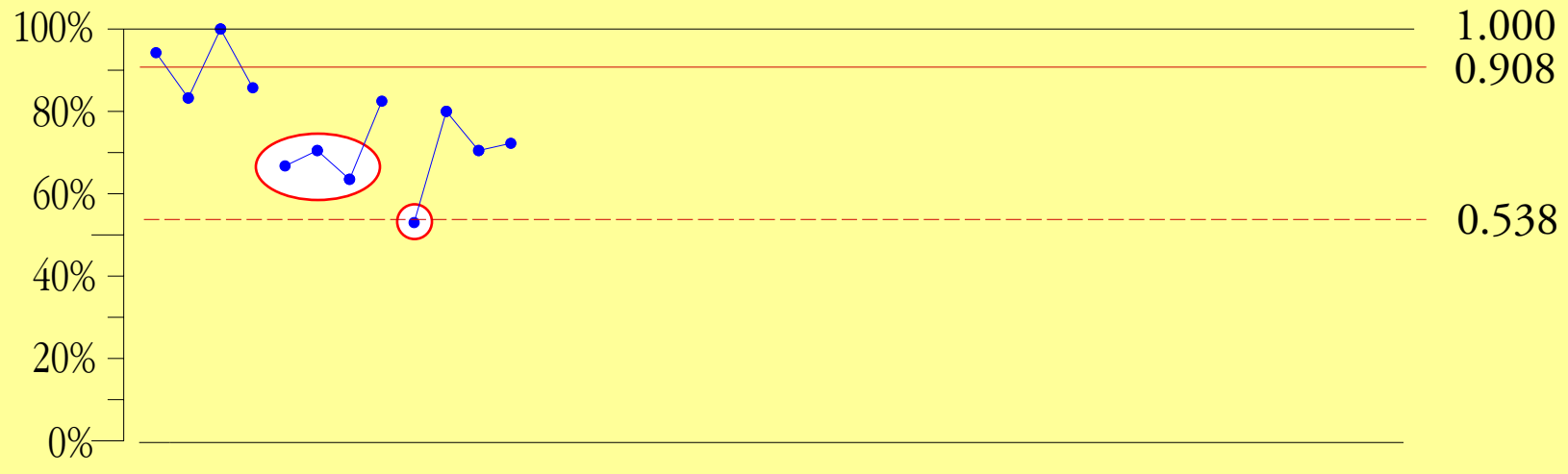
Year	1	2	3	4	5	6	7	8	9	10
Fatalities	38	60	67	57	49	46	53	45	32	40
Autopsies	35	44	46	32	36	24	16	16	13	26
Percent	92%	73%	69%	56%	73%	52%	57%	36%	41%	65%

These wide limits are only based on four data,

but they are tight enough

to allow us to begin to tell the story contained in these data.

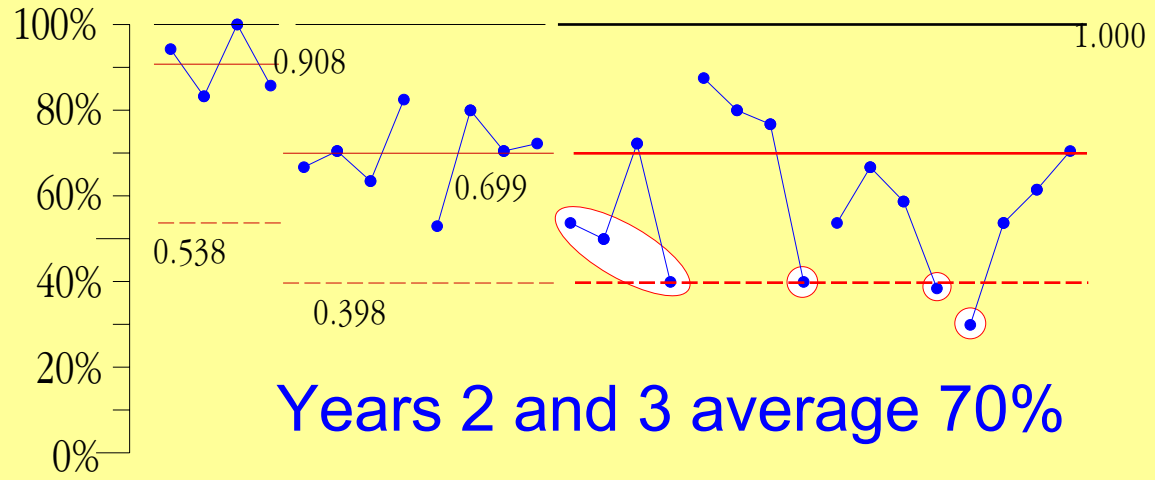
What Happens When the Data Are Full of Signals?



Year	1	2	3	4	5	6	7	8	9	10
Fatalities	38	60	67	57	49	46	53	45	32	40
Autopsies	35	44	46	32	36	24	16	16	13	26
Percent	92%	73%	69%	56%	73%	52%	57%	36%	41%	65%

With a run near the limit
and a point below the limit,
Years 2 and 3 are different from Year 1.
So we compute limits for Years 2 and 3.

What Happens When the Data Are Full of Signals?

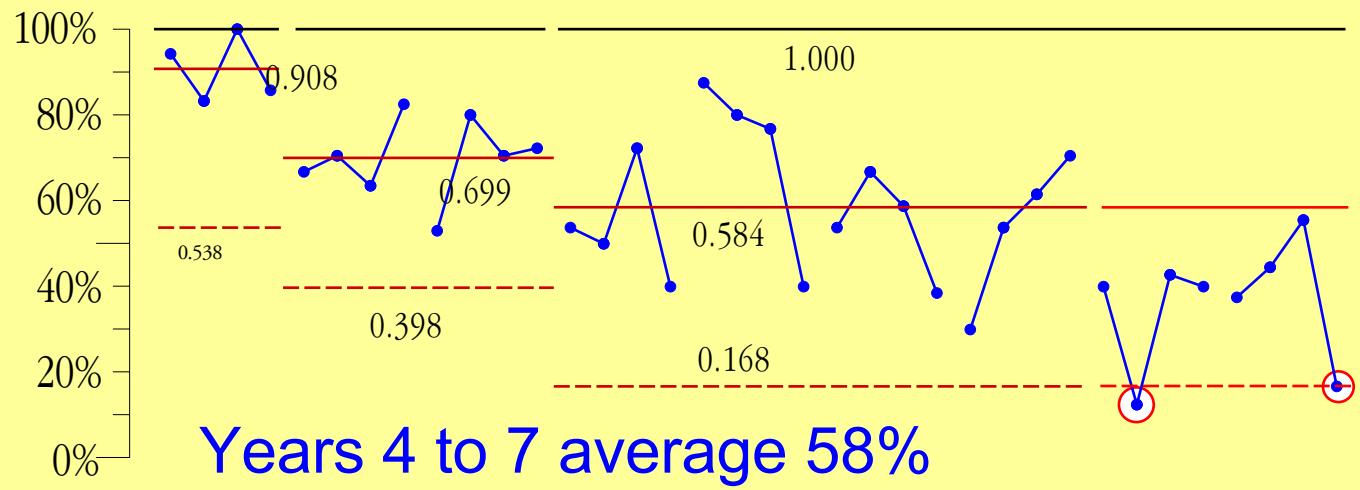


Years 2 and 3 average 70%

Year	1	2	3	4	5	6	7	8	9	10
Fatalities	38	60	67	57	49	46	53	45	32	40
Autopsies	35	44	46	32	36	24	16	16	13	26
Percent	92%	73%	69%	56%	73%	52%	57%	36%	41%	65%

With runs near the limit and points below the limit
Years 4, 5, 6, and 7
are detectably different from Years 2 and 3.
So we compute limits for Years 4 to 7.

What Happens When the Data Are Full of Signals?

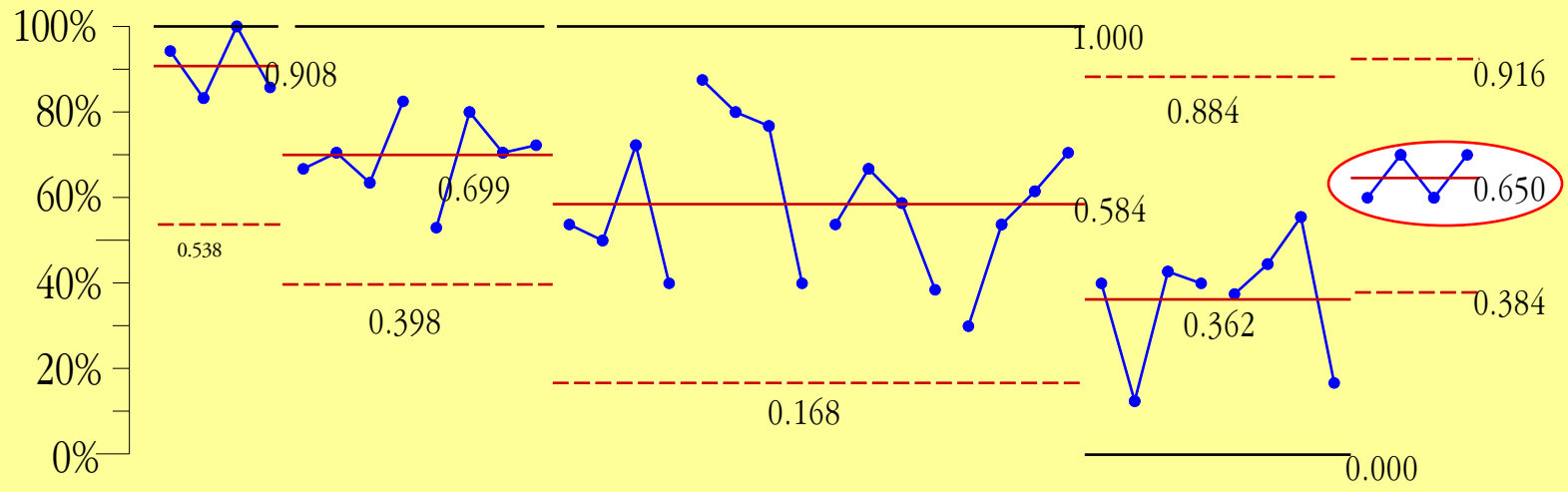


Years 4 to 7 average 58%

Year	1	2	3	4	5	6	7	8	9	10
Fatalities	38	60	67	57	49	46	53	45	32	40
Autopsies	35	44	46	32	36	24	16	16	13	26
Percent	92%	73%	69%	56%	73%	52%	57%	36%	41%	65%

With points below the limit,
Years 8 and 9
are detectably different from Years 4 to 7.

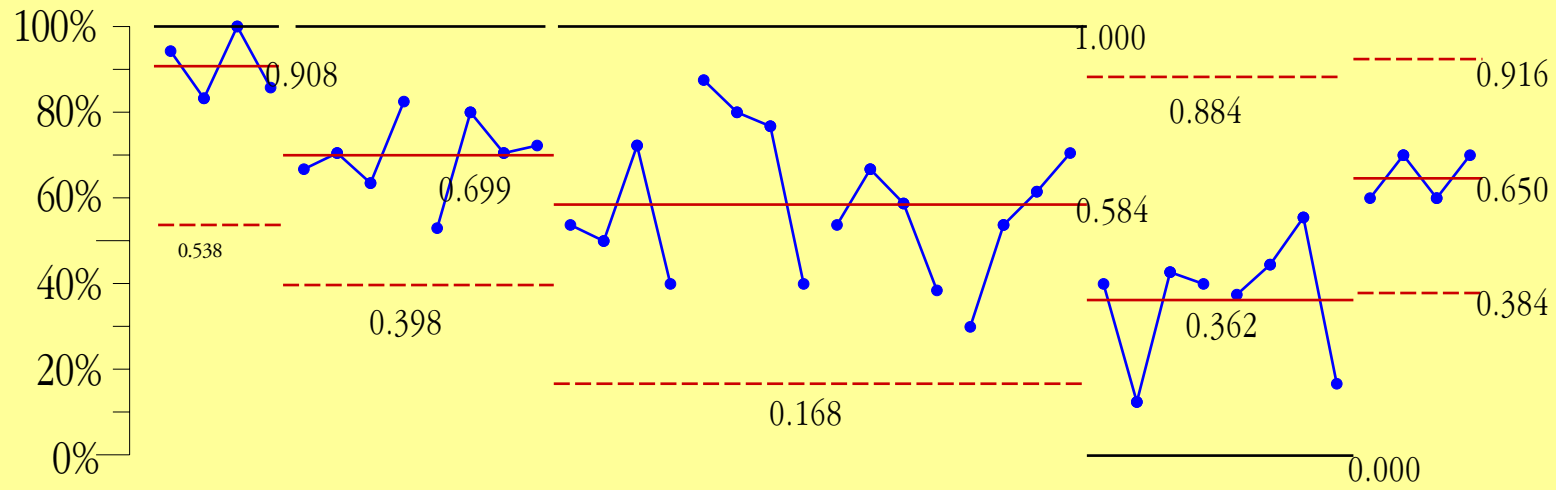
What Happens When the Data Are Full of Signals?



Year	1	2	3	4	5	6	7	8	9	10
Fatalities	38	60	67	57	49	46	53	45	32	40
Autopsies	35	44	46	32	36	24	16	16	13	26
Percent	92%	73%	69%	56%	73%	52%	57%	36%	41%	65%

Years 8 and 9 average 36%
With a run of four beyond one sigma,
Year 10 is detectably different from Years 8 and 9.
Thus, using the quarterly counts,
we find five different levels in these data.

What Happens When the Data Are Full of Signals?



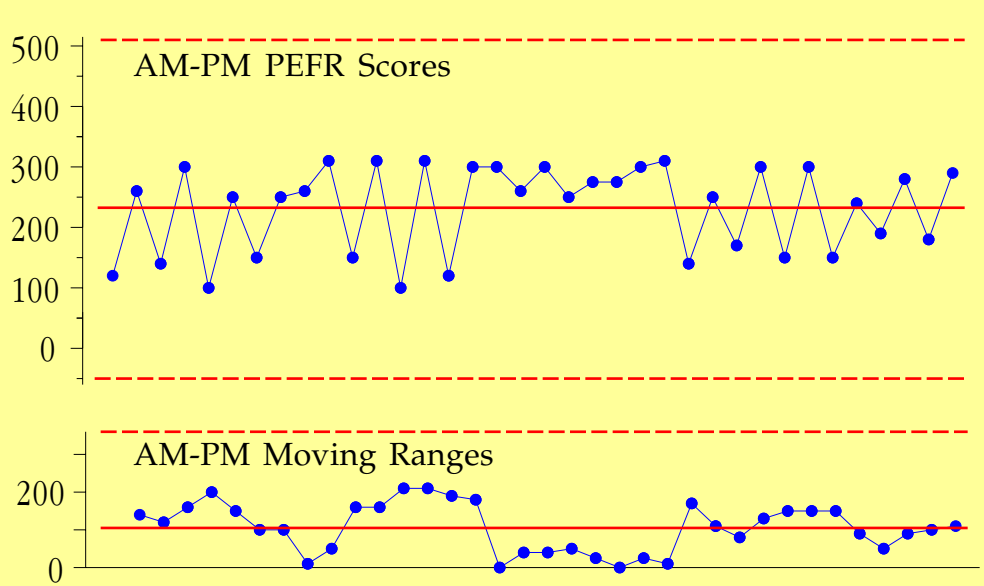
Year	1	2	3	4	5	6	7	8	9	10
Fatalities	38	60	67	57	49	46	53	45	32	40
Autopsies	35	44	46	32	36	24	16	16	13	26
Percent	92%	73%	69%	56%	73%	52%	57%	36%	41%	65%

In this hospital it was the job of the chaplain
to get permission for a neonatal autopsy.
There were five chaplains over this ten year period.
The changes occurred at the end of years 1, 3, 7, and 9.

When the Data Represent Two or More Processes

AM pre-medication and PM post-medication PEFR scores were plotted in their time order sequence.

Patient's predicted best PEFR score is 310.



514

232

–50

346

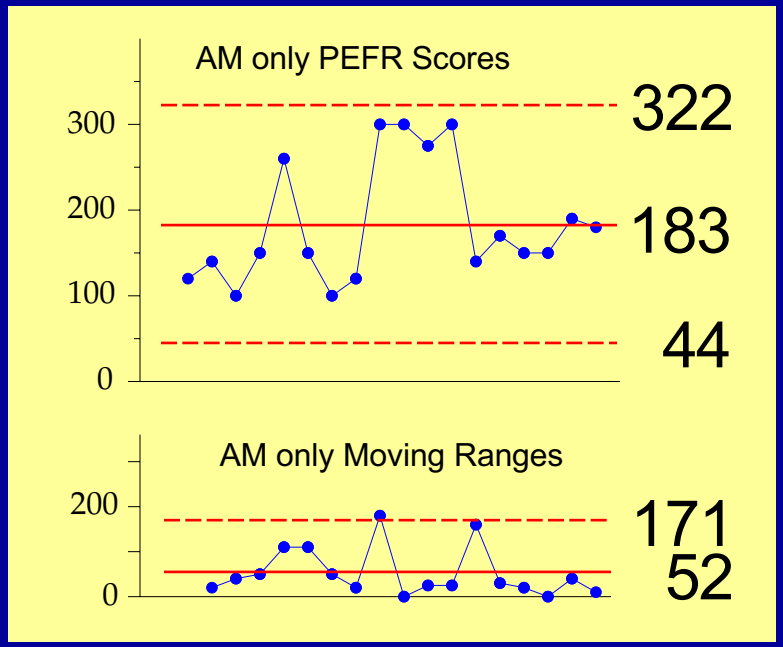
106

These limits make no sense

These Moving Ranges capture the daily swings rather than the routine variation from day-to-day.

When the Data Represent Two or More Processes

Charting only the AM pre-medication PEFR scores reveals the patient's chronic condition.



322 consistent with predicted best score.

44 consistent with patient's past medical history.

Moving ranges track day-to-day variation.

XmR charts have to compare apples to apples, not apples to oranges.

Wide Limits Happen

So when the data are full of noise,

when the data are full of signals,

or when the data represent two or more processes,

the secret to using the charts

to tell the story contained in the data

is to disaggregate the time series,

change the time periods,

and rationally organize the data

according to the principles given.

The Question of Homogeneity Is Primary

If the data are reasonably homogeneous,

then we are justified in assuming that

the underlying process is predictable.

Here we may use the central line, limits,

and other statistics computed from the data,

to characterize that underlying process.

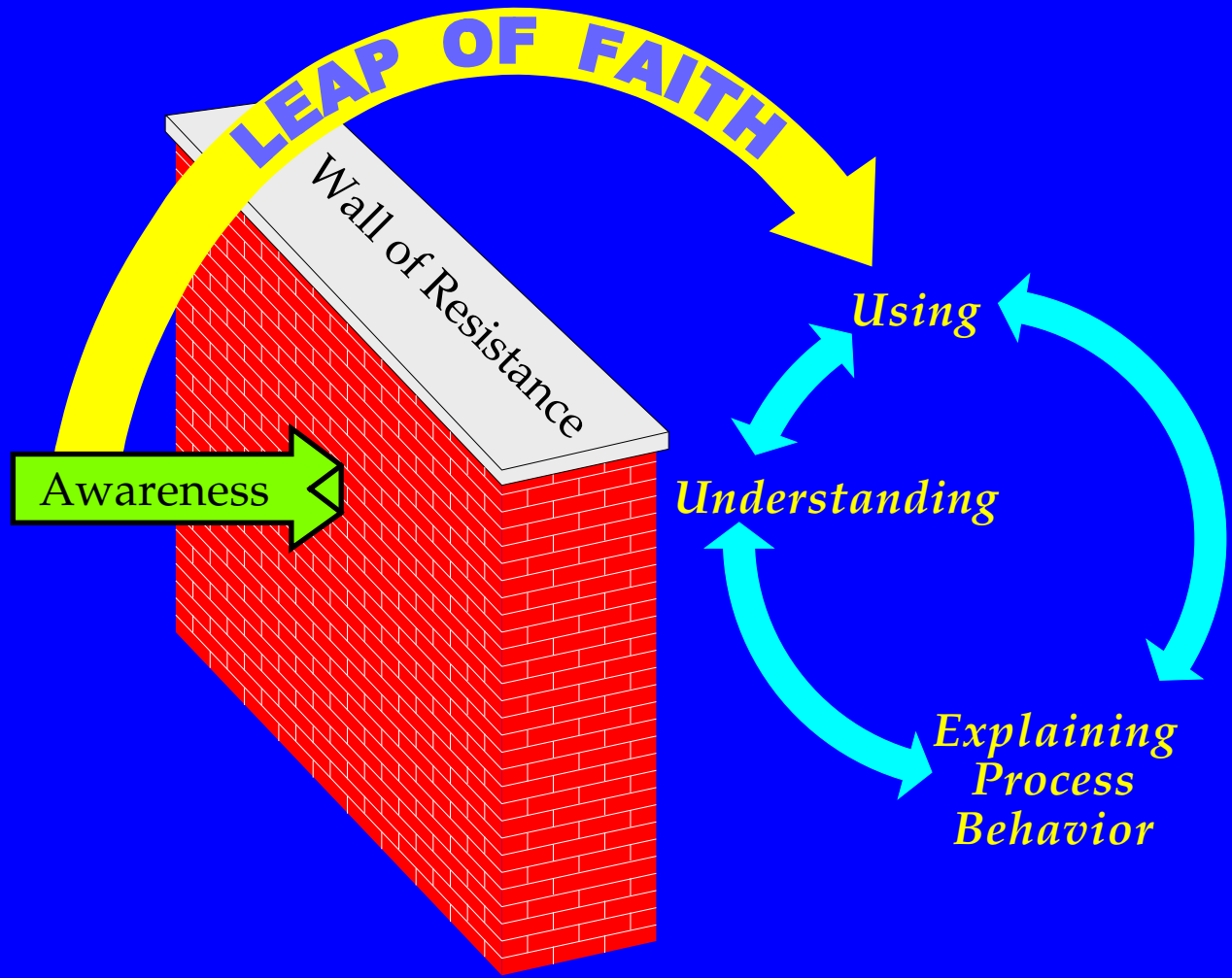
The Question of Homogeneity Is Primary

If, however, the data show that the process is changing,
the focus shifts from characterizing the process
to using the computed values to detect
when and how the process is changing.

Here there is no such thing as computing the right limits
or having enough data.

Until we understand the story told by the data
we will not be able to operate our processes
up to their full potential.

WHAT IT WILL TAKE



If you have questions: djwheeler@spcpress.com