Quality Monitoring in the Social Services, Part III: Data Analysis

Curtis McMillen, Ph.D.

We'll talk about…

Two kinds of data analysis that a quality monitoring professional might do:

- ***Looking at variation***
- ***Ways to examine trends over time (run charts, statistical process control charts).***
- ***Comparisons***

Deming's Lasting Legacy

Two Types of Knowledge

Subject Matter Knowledge

Subject Matter Knowledge: Knowledge basic to the things we do in life. Professional knowledge.

Profound Knowledge: The interplay of the theories of systems, variation, knowledge, and psychology.

Profound Knowledge

NICHQ
National Initiative for Children's Healthcare Quality

Employee Satisfaction – Greg's 7 Supervisees

6 items summed, each item scored 1-5, higher=better

29

28

28

27

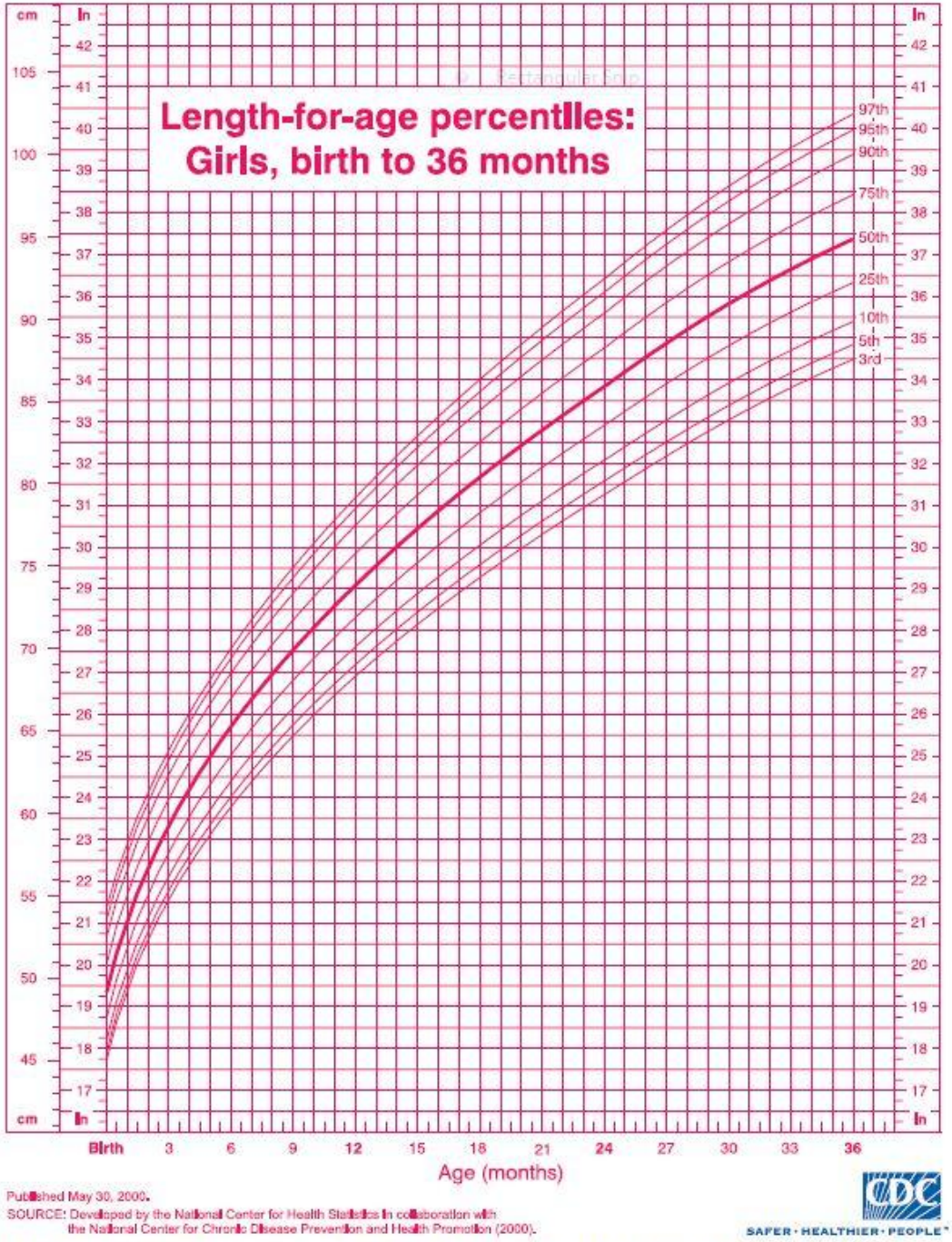
26

24

20

Variation

$\bar{X}=26$



gure 4. Individual growth chart 3rd, 5th, 10th, 25th, 50th, 75th, 90th, 95th, 97th percentiles, birth to 36 months: Girls length-for-age

Growth Chart – length – girls ages birth to 36 months, U.S.

Employee Satisfaction – Greg's 7 Supervisees

Score

29

28

28

27

26

24

20

$\bar{X}=26$

Variation

Employee Satisfaction – Greg's 7 Supervisees

Score	*substract mean*
29	-26 = 3
28	-26 = 2
28	*-26 = 2*
27	*-26 = 1*
26	*-26 = 0*
24	-26=-2
20	-26 = -6

$\bar{X}=26$

Variation

Employee Satisfaction – Greg's 7 Supervisees

Score	*substract mean*	*square it*
29	-26 = 3	9
28	-26 = 2	4
28	*-26 = 2*	*4*
27	*-26 = 1*	*1*
26	*-26 = 0*	*0*
24	-26=-2	4
20	-26 = -6	36

$\bar{X}=26$

Variation

Employee Satisfaction – Greg's 7 Supervisees

Score	*substract mean*	*square it*
29	-26 = 3	9
28	-26 = 2	4
28	-26 = 2	4
27	-26 = 1	1
26	-26 = 0	0
24	-26=-2	4
20	-26 = -6	36
$\bar{X}$=26		58=sum of squares

Variation

Employee Satisfaction – Greg's 7 Supervisees

Score	*substract mean*	*square it*
29	-26 = 3	9
28	-26 = 2	4
28	-26 = 2	4
27	-26 = 1	1
26	-26 = 0	0
24	-26=-2	4
20	-26 = -6	36
$\bar{X}$=26		58 / 6 df=9.67= variance

Sq rt of 9.67= 3.11= standard deviation

Variation

Variation is at the  of quality management

How much can we expect this indicator of quality to vary?

•How far does something have to vary from what is expected to indicate a problem?

How do our different programs / teams / employees vary in quality from one another?

•How far does something have to vary from what is expected to indicate a problem?

Do Consumer Satisfaction Scores Differ by Race?
Sunshine Outpatient Department: Consumer Satisfaction

30	28	30
28	25	30
27	24	27
26	24	26
26	24	25
26	23	24
26	23	24
26	22	24
25	21	23
25	20	22
25	20	21
24	20	20
24	19	18
23	18	14
23	17	
23	17	
22	17	
20	16	
19	15	
18	15	
17	11	
16	10	
13	8	
8	7	
6	6	
	6	
	6	

66 clients
Var= 41.51
SD= 6.44
Mean = 20.24
Total Sum of squares=2698

Do Consumer Satisfaction Scores Differ by Race?
Sunshine Outpatient Department: Consumer Satisfaction

White Clients	*Afric Am Clients*	*Latina/a Clients*
30	28	30
28	25	30
27	24	27
26	24	26
26	24	25
26	23	24
26	23	24
26	22	24
25	21	23
25	20	22
25	20	21
24	20	20
24	19	18
23	18	14
23	17	
23	17	
22	17	
20	16	
19	15	
18	15	
17	11	
16	10	
13	8	
8	7	
6	6	
	6	
	6	

Do Consumer Satisfaction Scores Differ by Race?
Sunshine Outpatient Department: Consumer Satisfaction

White Clients

30
28
27
26
26
26
26
26
25
25
25
24
24
23
23
23
22
20
19
18
17
16
13
8
6

Mean= 21.84
Var= 36.06
SD= 6.00
Sum of squares= 866

Afric Am Clients

28
25
24
24
24
23
23
22
21
20
20
20
19
18
17
17
17
16
15
15
11
10
8
7
6
6
6

Mean= 17.11
Var= 42.88
SD= 6.55
Sum of squares= 1114.6

Latina/a Clients

30
30
27
26
25
24
24
24
23
22
21
20
18
14

Mean= 23.43
Var= 19.03
SD= 4.36
Sum of squares= 247.7

Do Consumer Satisfaction Scores Differ by Race?
Sunshine Outpatient Department: Consumer Satisfaction

White Clients

30 28 27 26 26 26 26 26 25 25 25 24 24 23 23 23 22 20 19 18 17 16 13 8 6

Mean= 21.84
Var= 36.06
SD= 6.00
Sum of squares= 866

Within groups sum of squares= 866+ 1115 + 248=2228

Mean within group sum of squares = wss/df= 2228/63=35.35

Afric Am Clients

28 25 24 24 24 23 23 22 21 20 20 20 19 18 17 17 17 16 15 15 11 10 8 7 6 6 6

Mean= 17.11
Var= 42.88
SD= 6.55
Sum of squares= 1114.6

Latina/a Clients

30 30 27 26 25 24 24 24 23 22 21 20 18 14

Mean= 23.43
Var= 19.03
SD= 4.36
Sum of squares= 247.7

Do Consumer Satisfaction Scores Differ by Race?
Sunshine Outpatient Department: Consumer Satisfaction

30	28	30
28	25	30
27	24	27
26	24	26
26	24	25
26	23	24
26	23	24
26	22	24
25	21	23
25	20	22
25	20	21
24	20	20
24	19	18
23	18	14
23	17	
23	17	
22	17	
20	16	
19	15	
18	15	
17	11	
16	10	
13	8	
8	7	
6	6	
	6	
	6	

66 clients
Var= 41.51
SD= 6.44
Mean = 20.24
Total Sum of squares=2698

Between groups sum of squares=

$$n_1(\bar{x}_1 - \bar{x})\text{^}2 + n_2(\bar{x}_2 - \bar{x})\text{^}2 + n_3(\bar{x}_3 - \bar{x})\text{^}2$$

Mean sum of squares is this number divided by df

470/2 = 235

Mean between group sum of squares /
mean within group sum of squares = *f statistic*

Hypothesis Testing

Null hypothesis: Consumer satisfaction scores are the same for each ethnic group

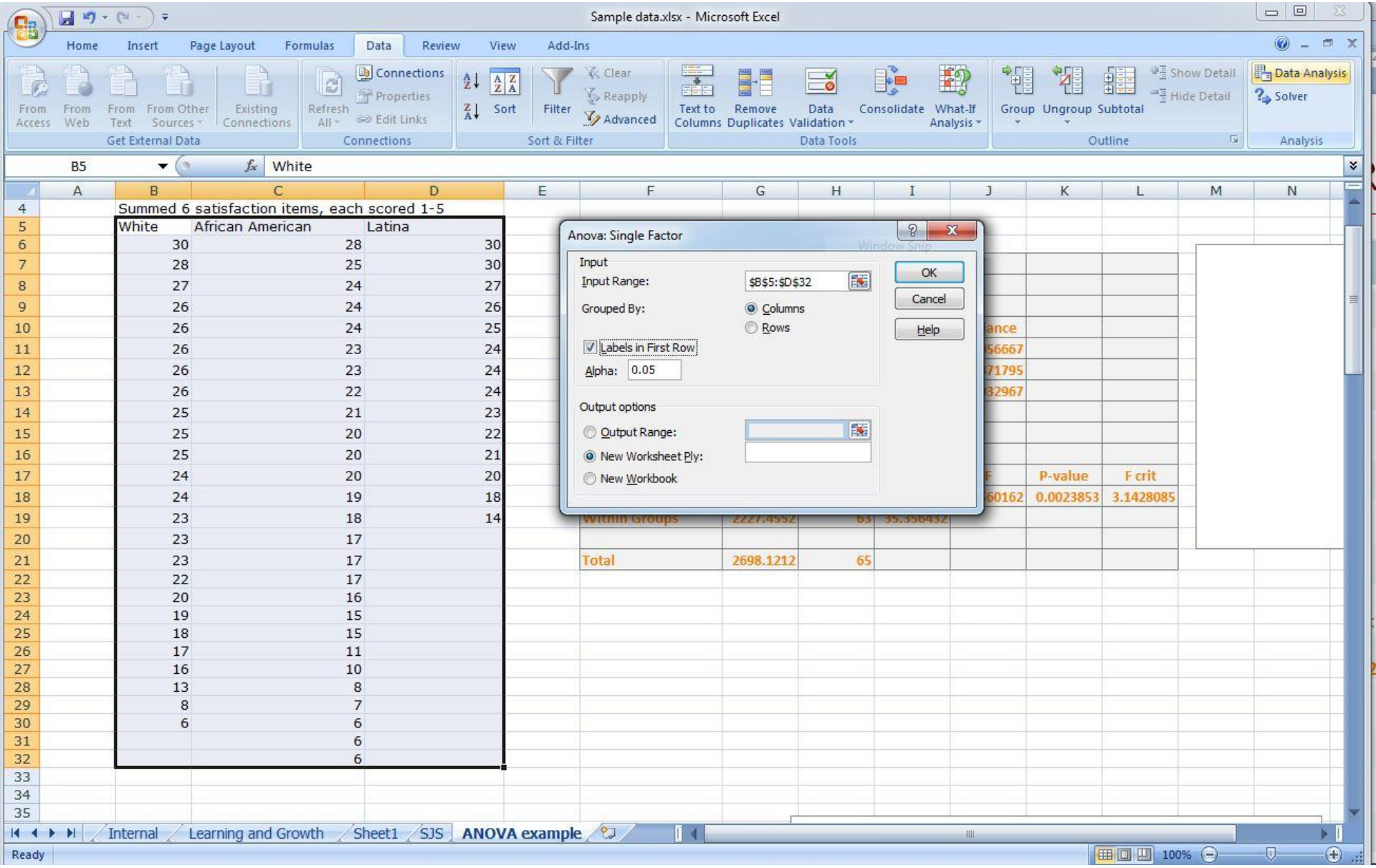
Mean for whites=mean for African Americans=mean for Latino/as

Alternative hypothesis: Consumer satisfaction scores differ by ethnicity.

You see if there is enough evidence to reject the null hypothesis.

Do Consumer Satisfaction Scores Differ by Race?

Anova: Single Factor						
SUMMARY						
Groups	Count	Sum	Average	Variance		
White	25	546	21.84	36.056667		
African American	27	462	17.111111	42.871795		
Latina	14	328	23.428571	19.032967		
ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	470.66597	2	235.33299	6.6560162	0.0023853	3.1428085
Within Groups	2227.4552	63	35.356432			
Total	2698.1212	65				


Sample data.xlsx - Microsoft Excel
Summed 6 satisfaction items, each scored 1-5
White
African American
Latina
Anova: Single Factor
Input
Input Range: $B$5:$D$32
Grouped By: Columns Rows
Labels in First Row
Alpha: 0.05
Output options
Output Range:
New Worksheet Ply:
New Workbook
OK
Cancel
Help
P-value
F crit
0.0023853
3.1428085
Total
2698.1212
65
Internal
Learning and Growth
Sheet1
SJS
ANOVA example
Ready

Let's do this by hand. Do the satisfaction scores among our four supervisors differ?

	Bob	Bonnie	Betty	Bernie
	29	28	25	22
	28	28	24	22
	28	27	23	21
	27	27	23	21
	26	25	23	21
	26	25	23	21
	24	20	23	20
	20	20	20	20
mean	26	25	23	21

GRAND MEAN=23.75
TOTAL SUM OF SQUARES= 270
Between groups degrees of freedom=3
Within groups degrees of freedom=28
WHAT ELSE DO YOU NEED TO CALCULATE?

The F statistic

Is a ratio of one number to another.

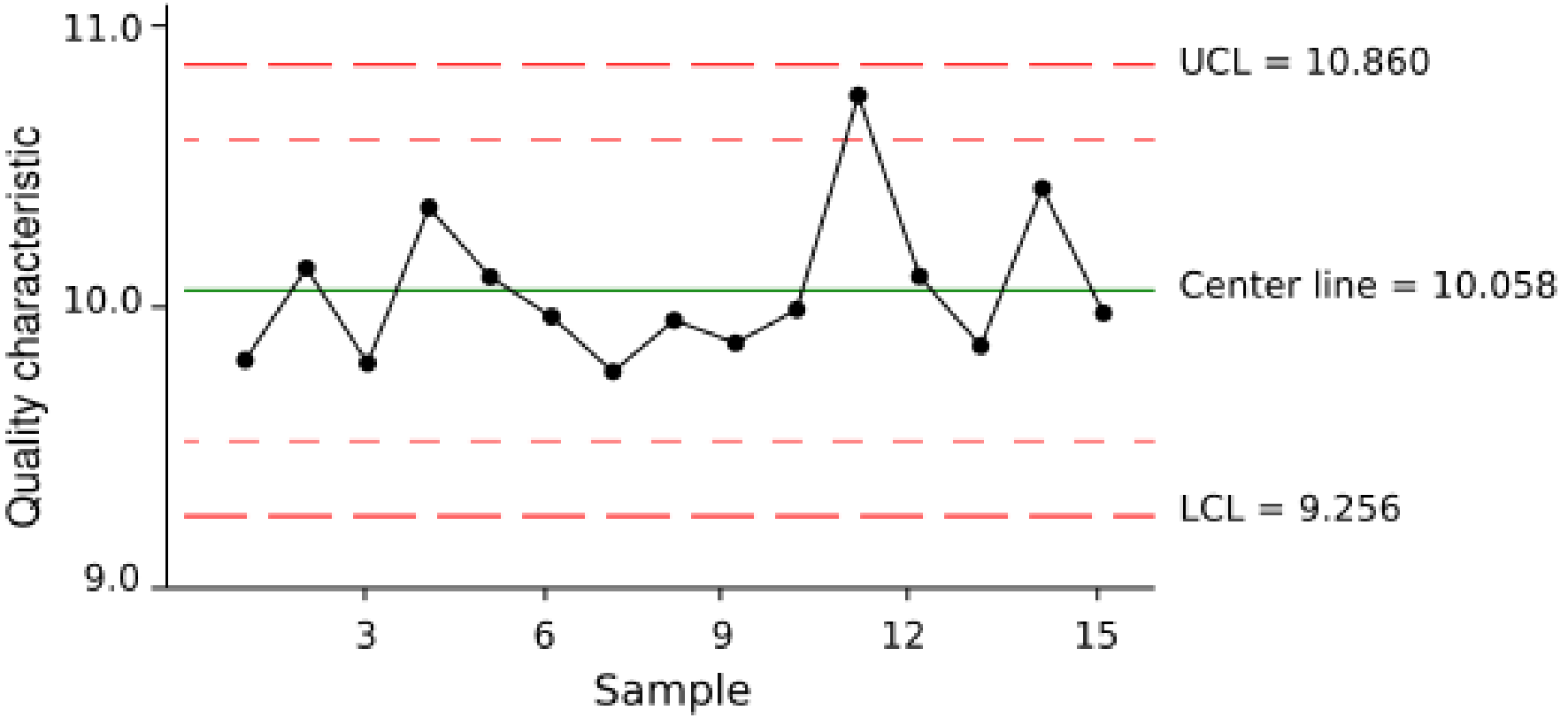
The higher the ratio, the greater the likelihood that your scores differ across groups.

The critical value – the point at which you say you can reject the null varies by the number of degrees of freedom.

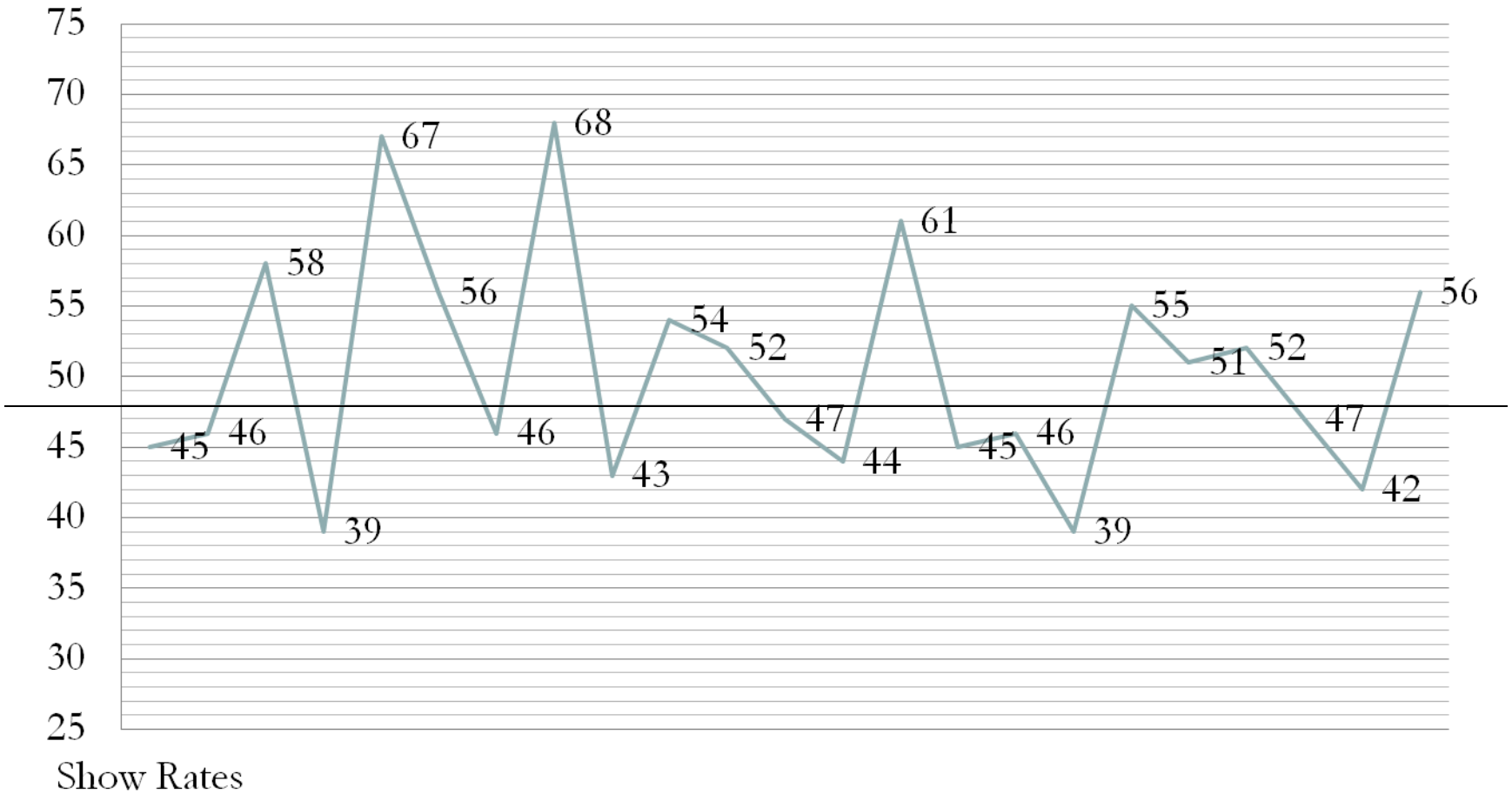
http://www.danielsoper.com/statcalc3/calc.aspx?id=4

Anova: Single Factor						
SUMMARY						
Groups	Count	Sum	Average	Variance		
Bob	8	208	26	8.285714 3		
Bonnie	8	200	25	10.85714 3		
Betty	8	184	23	2		
Bernie	8	168	21	0.571428 6		
ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	118	3	39.33333 3	7.245614	0.000958 2	2.946685 3
Within Groups	152	28	5.428571 4			
Total	270	31				

QualityTools: Run and Statistical Control Charts

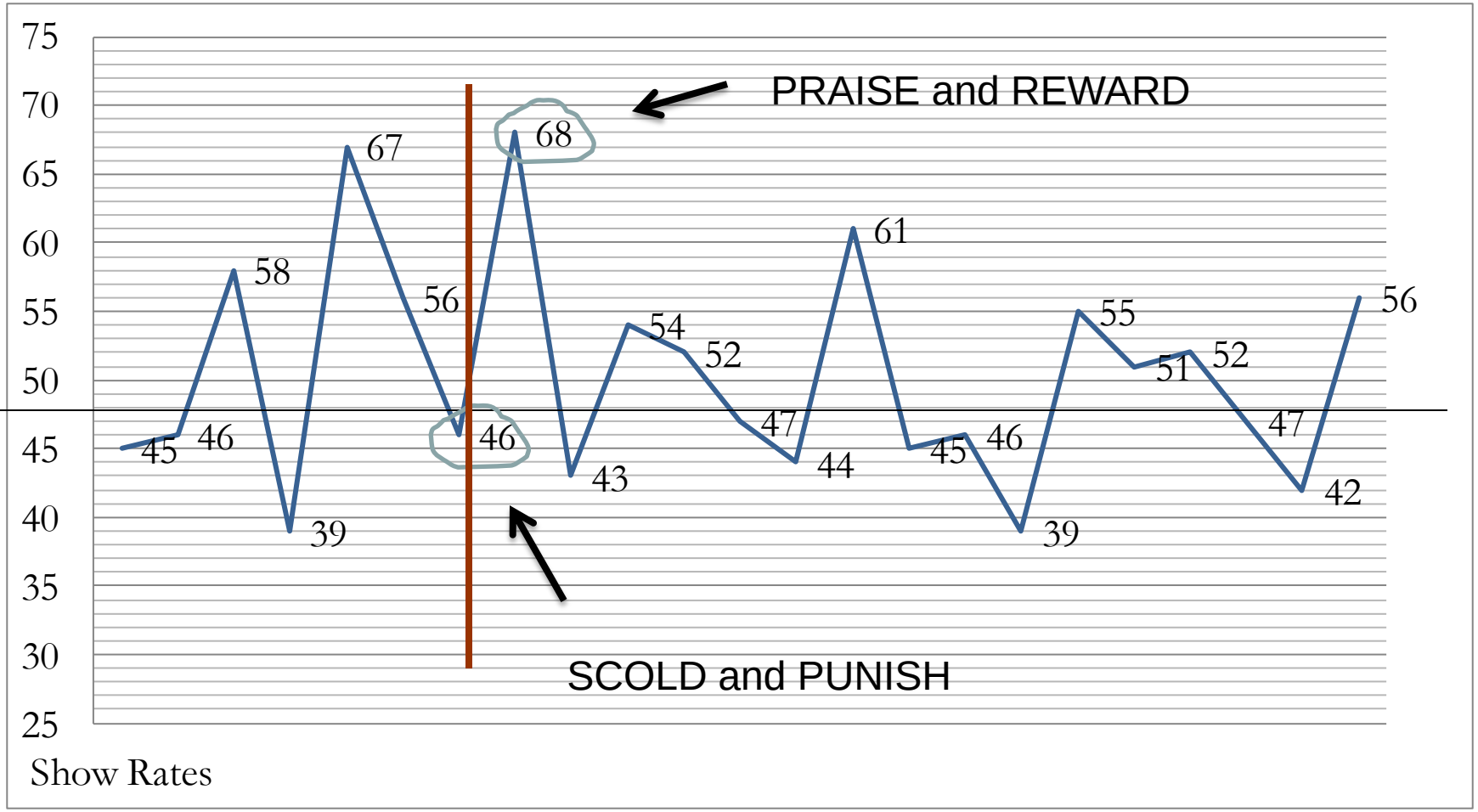


A run chart for counseling appointment show rates over 24 months



Mean=48%; median=47%

A run chart for counseling appointment show rates over 24 months



Mean=48%; median=47%

Trend spotting dangers

- *People will see trends where there are none.*
- *People will try to explain natural variation as special events.*
- *People will blame and give credit for things over which they have no control.*

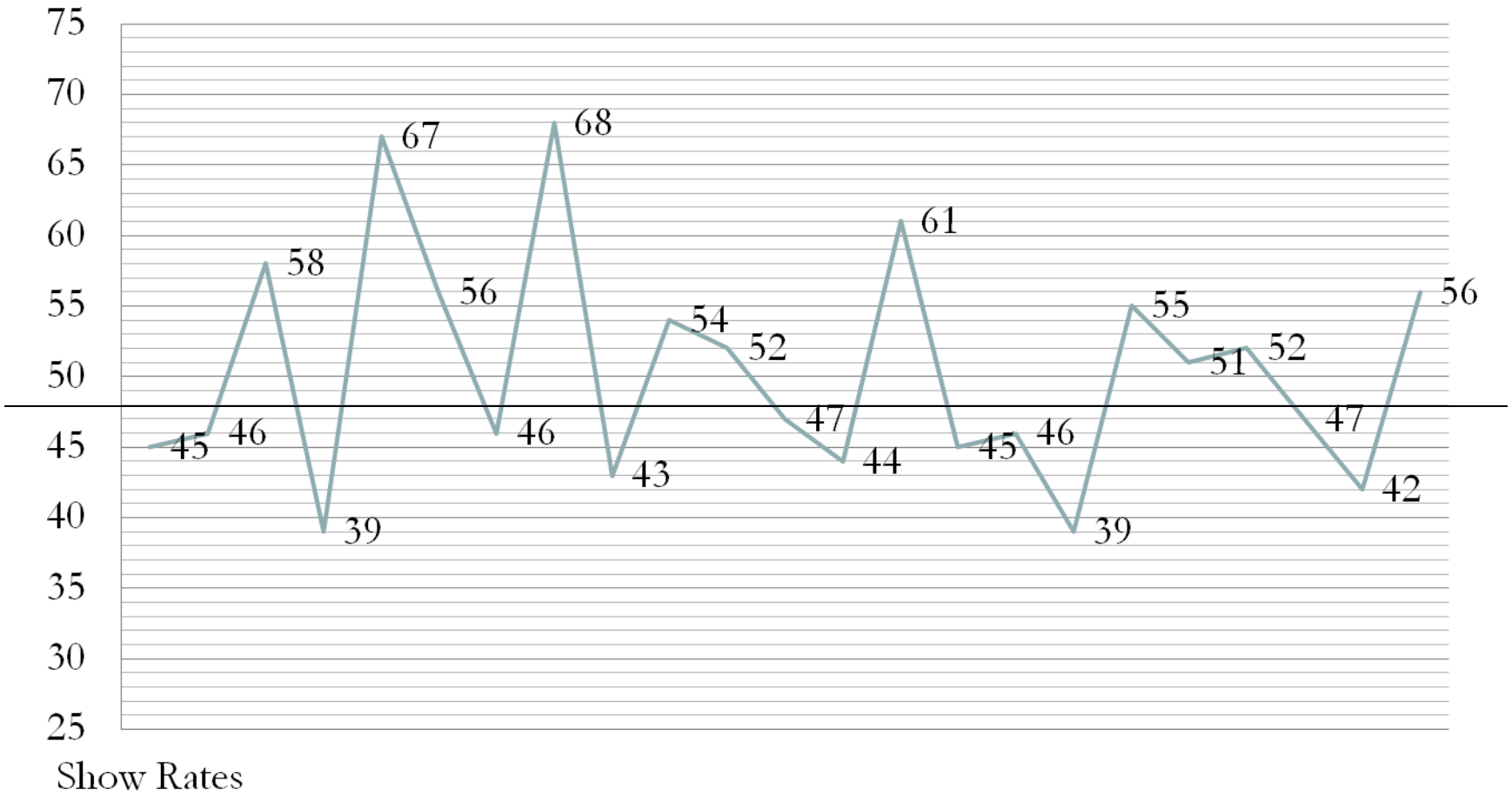
Lloyd, R. (2004). Quality health care: A guide to developing and using indicators. Sudbury, MA: Jones and Bartlett.

Variation

- ***Common Cause Variation*** = *the amount of variation that is natural or expected given the phenomenon of interest.*

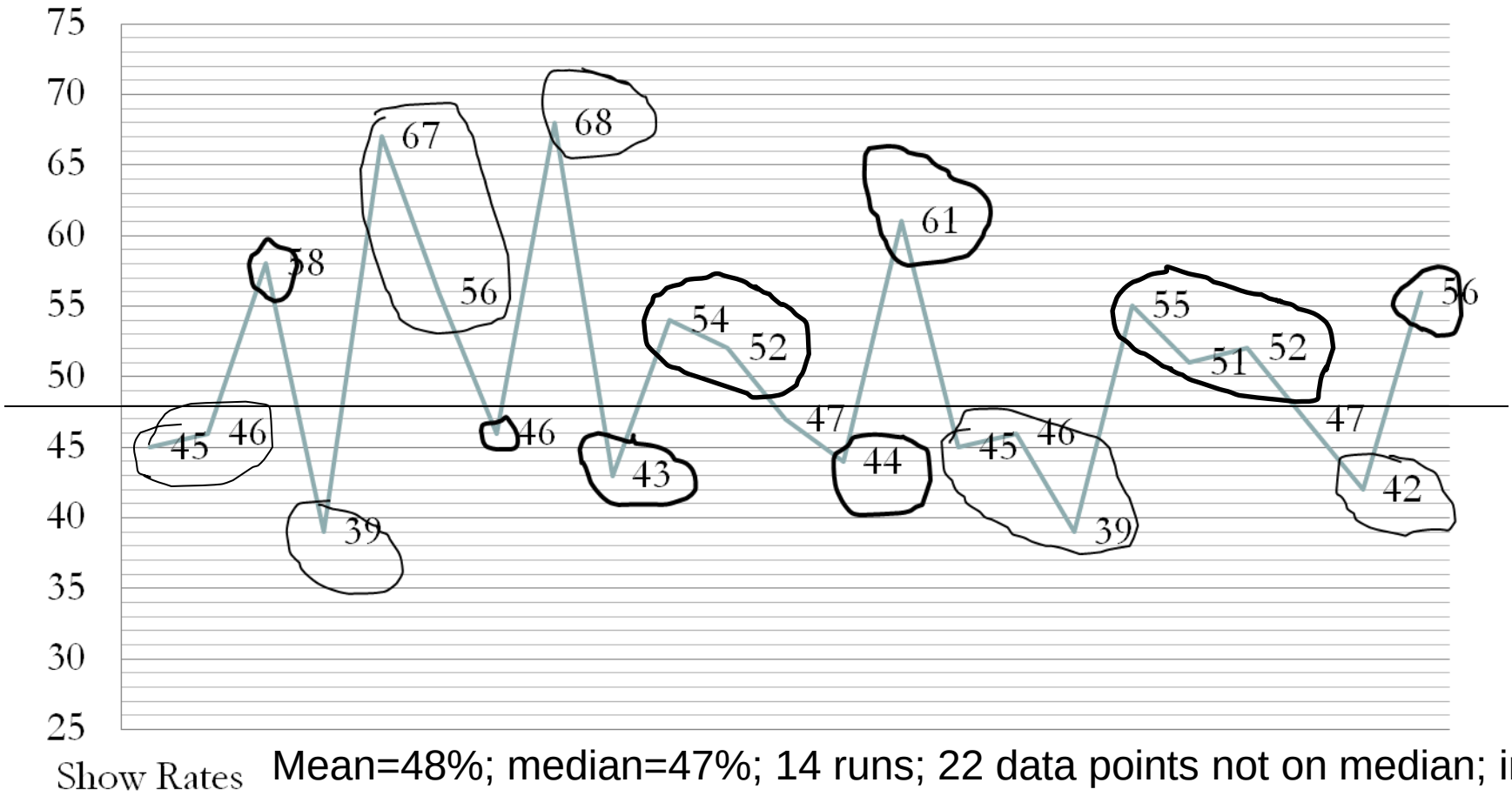
- ***Special Cause Variation*** = *a conclusion that the variation observed is due to something other than common cause variation.*

Is it normal variation or special cause variation?



Mean=48%; median=47%

A run chart for counseling appointment show rates over 24 months

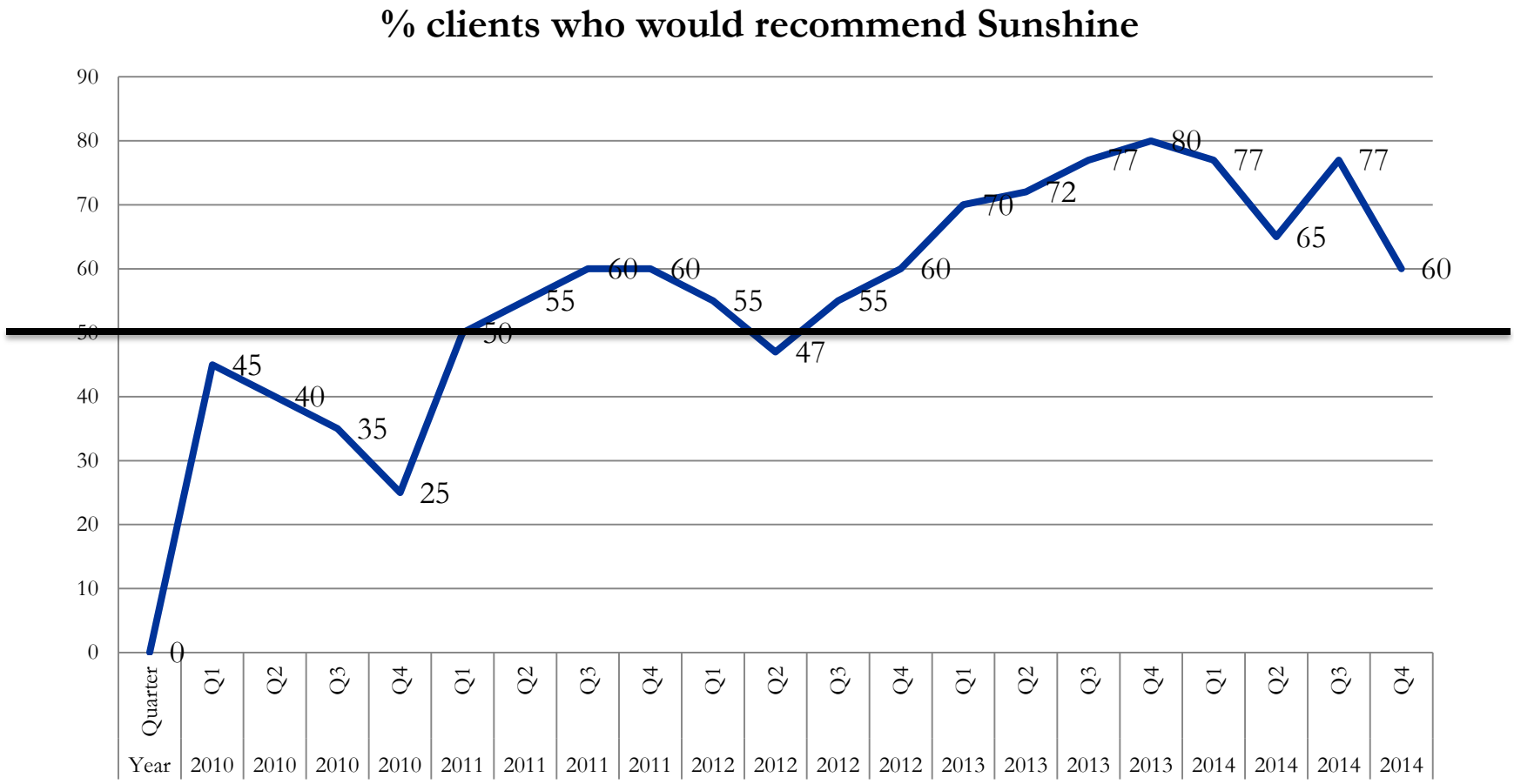


Mean=48%; median=47%; 14 runs; 22 data points not on median; in our case, statisticians would say we have special cause variation if we have 6 or fewer runs or 16 or more runs. We have no special cause variation indicated here.

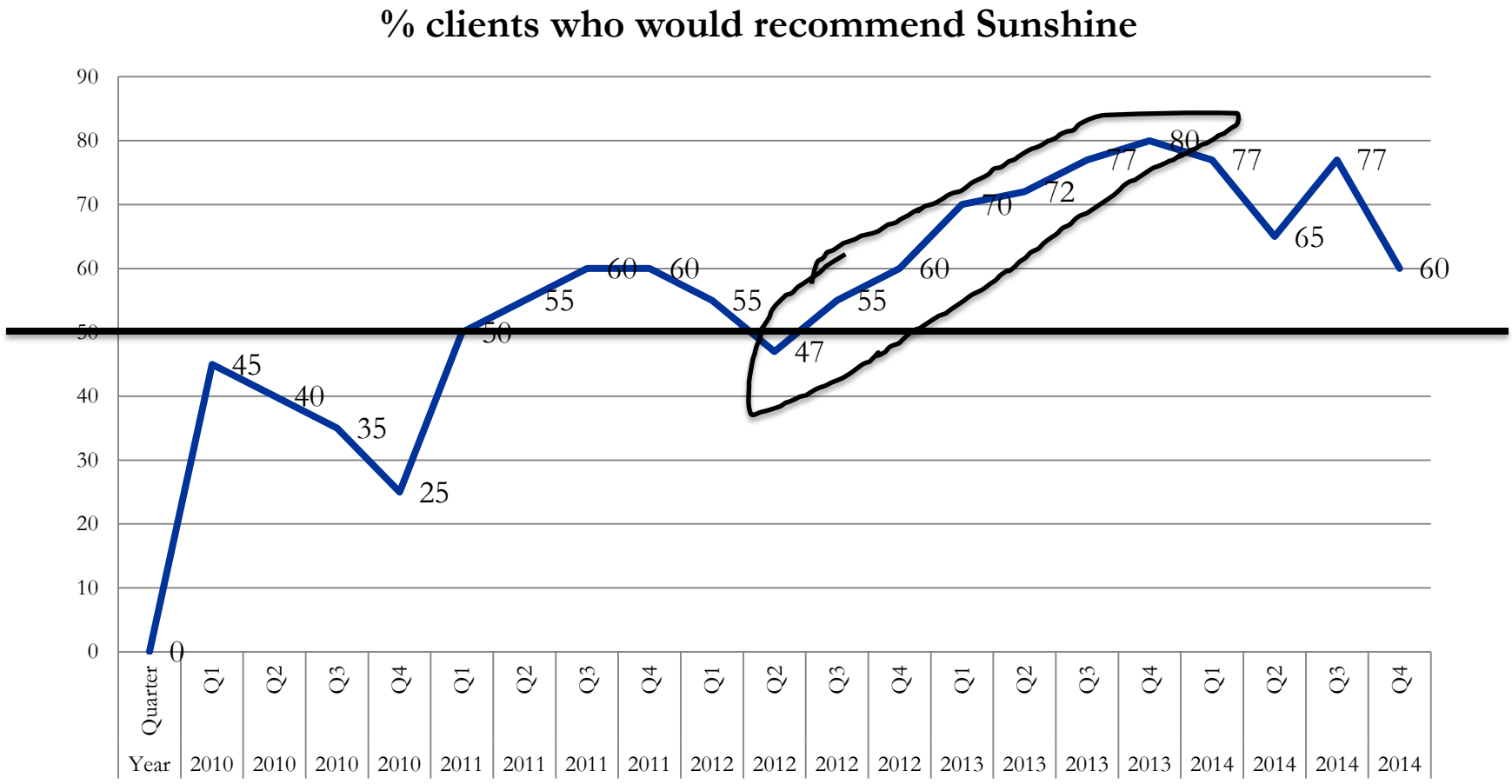
Rules for determining special cause variation using a run chart

- *A run of 8 in a row, 10 of 11, 12 of 14, 16 of 20 on the same side above the center line. (Some require 9 in a row on the same side of the center line).*
- *If you are looking to see if you have shifted an outcome based on a change in process that was instituted, the rule of thumb is that you need these same numbers (8 in row, 10 of 11, etc). Assumes you started with 21 or more data points.*
- *Six consecutive data points increasing or decreasing.*

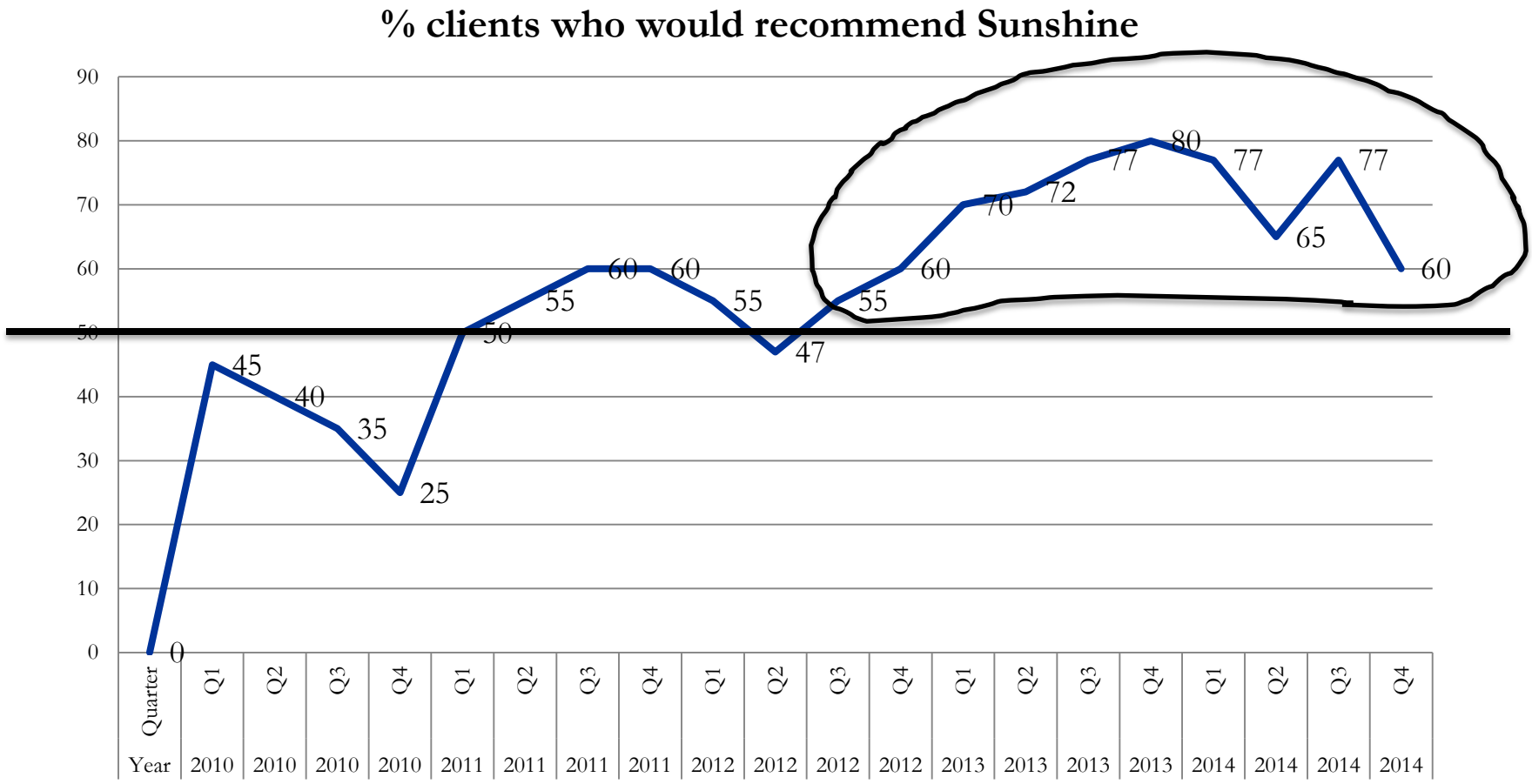
Interpret this run chart



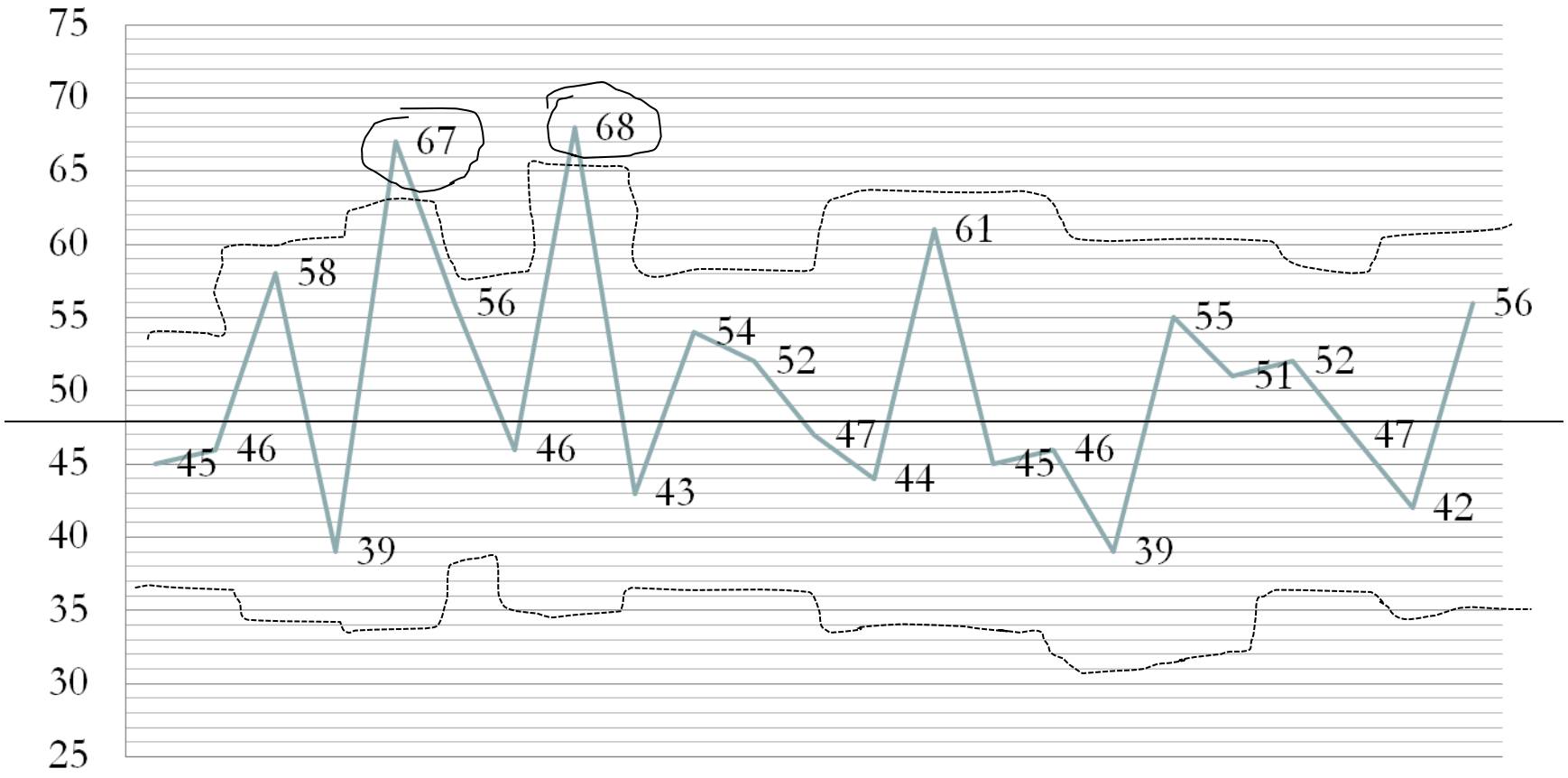
Interpret this run chart



Interpret this run chart

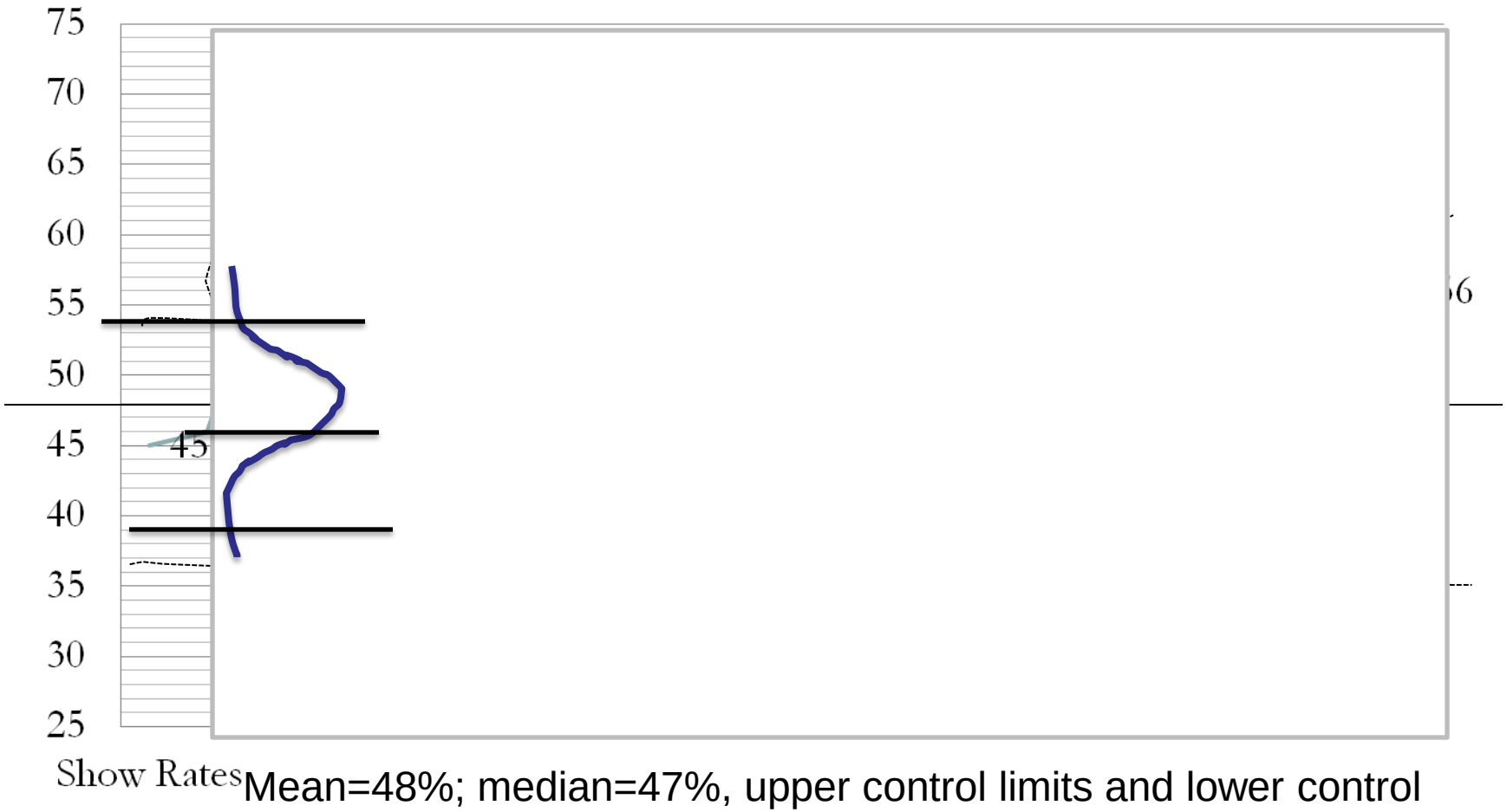


A statistical control p chart for counseling appointment show rates over 24 months



Mean=48%; median=47%, upper control limits and lower control limits differ because # of appts. differs each month. Statistics suggest two months might have special cause variation.

A statistical control p chart for counseling appointment show rates over 24 months

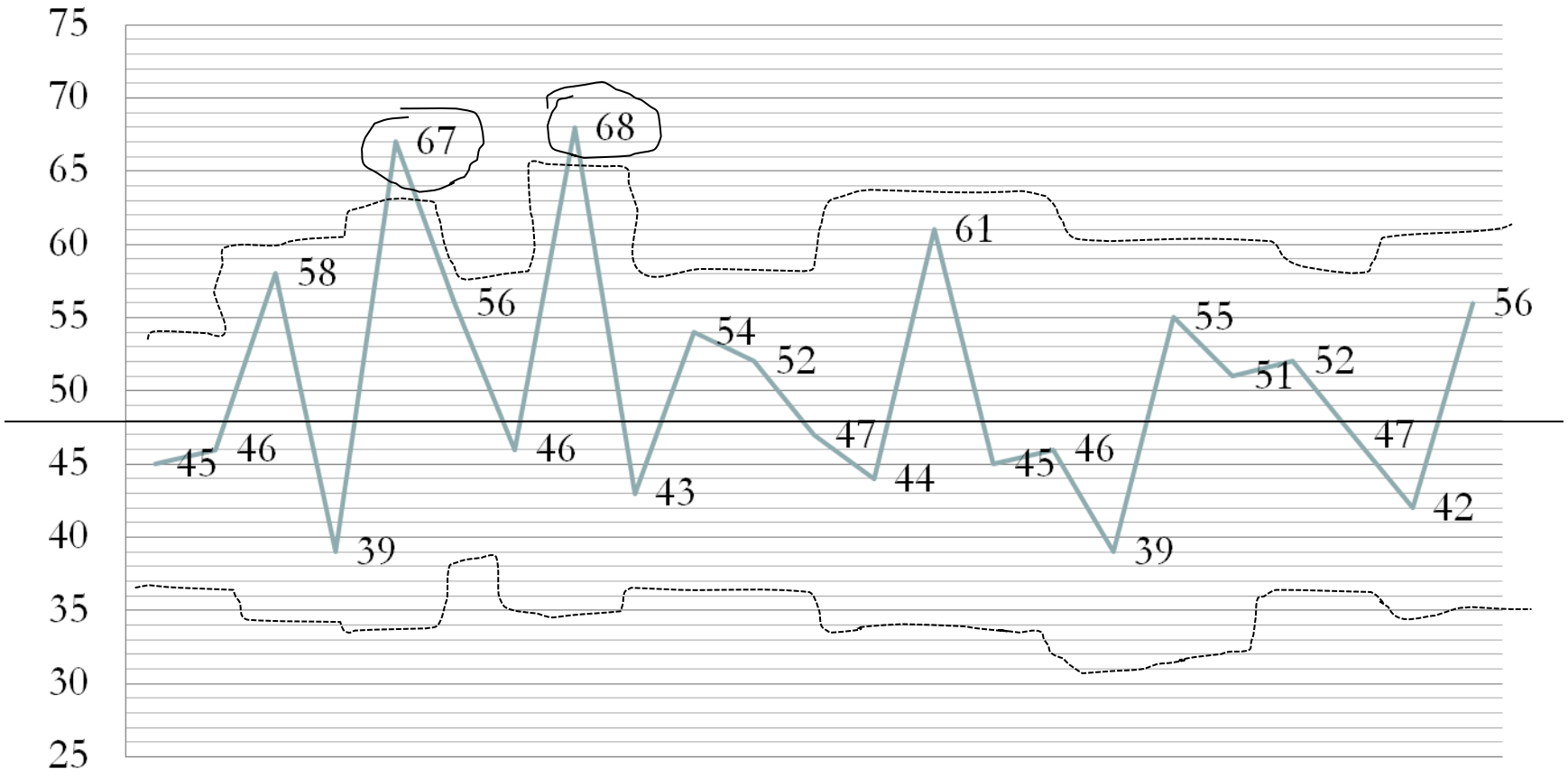


Mean=48%; median=47%, upper control limits and lower control limits differ because # of appts. differs each month. Statistics suggest two months might have special cause variation.

A statistical control p chart for counseling appointment show rates over 24 months

Mean=48%; median=47%, upper control limits and lower control limits differ because # of appts. differs each month. Statistics suggest two months might have special cause variation.

A statistical control p chart for counseling appointment show rates over 24 months



Mean=48%; median=47%, upper control limits and lower control limits differ because # of appts. differs each month. Statistics suggest two months might have special cause variation.

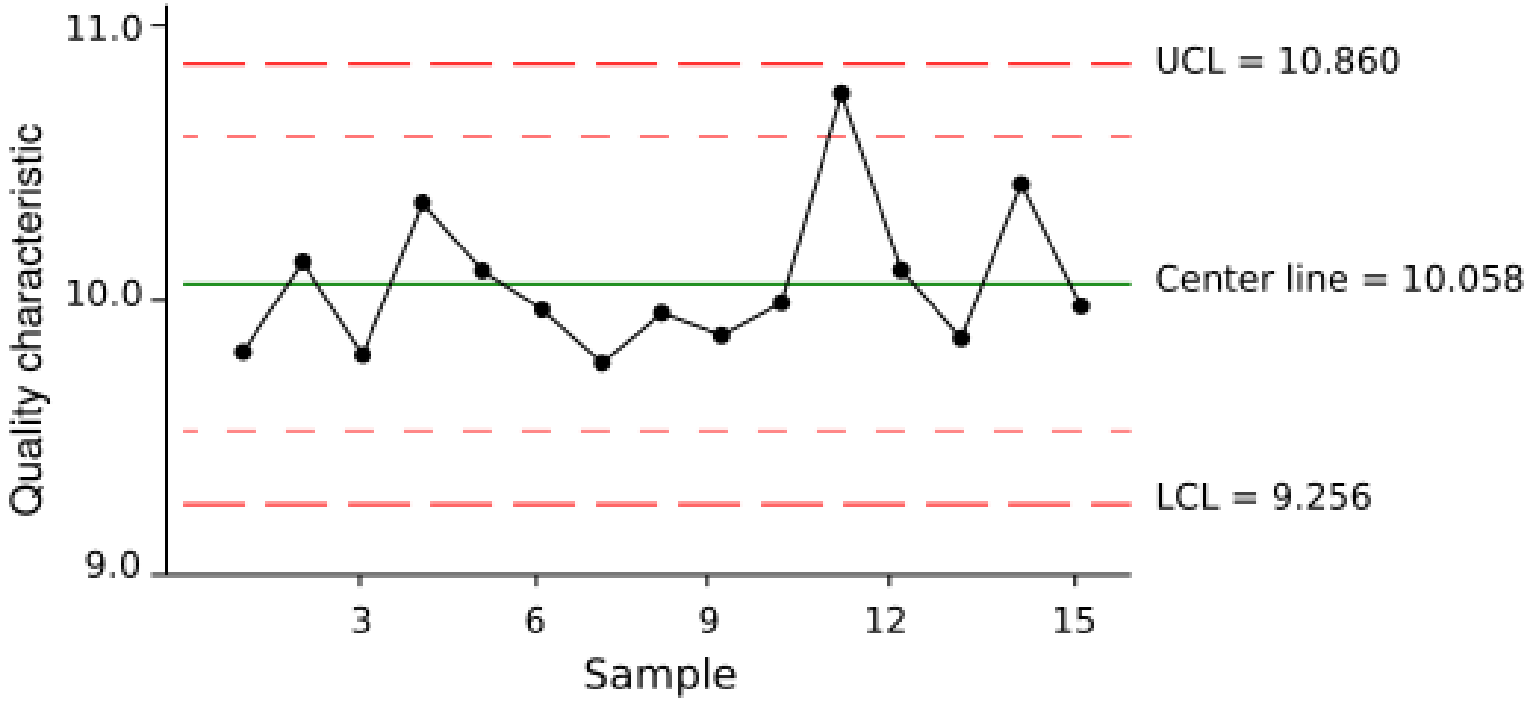
Types of statistical control charts

Name	Description
p chart	Monitors the percentage of something over time, when each thing of interest can have the characteristic or not. An example is a show rate. Think proportion.
$\bar{X}$ chart	(pronounced x-bar chart) Monitors the average of a sample over time. Example: mean consumer satisfaction.
Np chart	Monitors the number of something (historically, the number of nonconforming units in a sample). An example could be the number of nursing home patients who experienced a fall in the past month. Only one possible event per person.
u chart	Also measures number, but the number can vary per unit (3 of errors, each product can have more than one error). Example: number of falls.
c chart	Monitors the count of something, but the number can vary. Requires fixed sample size. Use Poisson distribution. Example: number of noncompliance errors found in a monthly audit of 100 charts.

Each type of chart

has a different <u>formula</u> for calculating

- **the center line**
- **the upper limit**
- **The lower limit**



Comparison to a Benchmark

Benchmarking = process of comparing one organizations performance metrics against (known best practice metrics, industry bests, bests in other industries).

Benchmarking processes may include visiting the high performing organizations to see how they managed to get their metrics so high/ low.

Where do I get a benchmark?

- *For number of restraint incidents in an adolescent psychiatric unit?*

Where do I get a benchmark?

- *For number of restraint incidents in an adolescent psychiatric unit?*

A state association? A state credentialing unit? The literature.

39 restraints per 1000 patient days

LeBel et al. (2004). Child and adolescent inpatient restraint reduction: A state initiattive to promote strength-based care. Journal of the American Academy of Child and Adolescent Psychiatry, 43, 37-45.

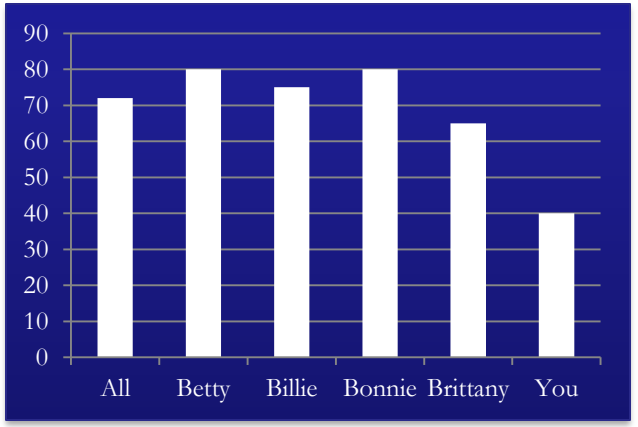
Giving Negative Feedback

- *"Your unit scores are the worst in the agency."*
- *"There were more deficiencies cited in your charts than in any one else's."*
- *"You suck at this."*

Precise, neutral tone, avoiding judgmental language

Giving Negative Feedback

- *INSTEAD OF: "Your unit scores are the worst in the agency."*
- *"Here are the scores for the agency and of all the teams. This one is your team."*



Precise, timely, neutral tone, avoiding judgmental language

Giving negative feedback

- *Is this what you had expected.*
- *If you had to guess, what are behind these scores?*
- *When can you move to problem solving?*