

Level 5 – Continuous Improvement with Industry 4.0

Measure 3 – Variation and Statistical Process Control (SPC)



Measure - Where are we?

Define	Measure	Analyse	Improve	Control
✓ Introduction	✓ Measurement	☐ Root Cause	☐ PDCA	☐ Auditing
✓ $Y=f(x)$	✓ MSA	☐ Regression	☐ Risk	☐ Lessons Learnt
✓ SIPOC	✓ Data Collection	☐ Hypothesis	Management	☐ SOPs
✓ Pareto	☐ Variation	Testing	☐ FMEA	☐ SMED
✓ Cause & Effect	☐ SPC	☐ t-tests	☐ Value Stream	☐ TPM
✓ 5 Whys	☐ Process	☐ ANOVA	Mapping	☐ Control Case
✓ Is it? Is it not?	Capability	☐ DoE	☐ Solution	Study
✓ Project Charter	☐ Charting		Selection	
✓ VoC & QFD	☐ OEE		Matrix	
✓ Team Working			☐ Five S	
✓ Critical to			☐ Benchmarking	
Quality (CTQ)			☐ Control Plans	



Measure - Data Collection Plan

MADE

1. What data will be collected?
2. Why is it needed?
3. Who is responsible?
4. How will it be collected?
5. When will it be collected?
6. Where will it be collected?



Develop your
Data Collection
Plan around
these questions.



What data will be collected?

Will you be measuring inputs (x) to a process or outputs (y)

	Y	X1	X2	X3
Existing data				
New Data				



Data Collection Sheets – why?

- **Standardize the data collection process**
- **Organises the data into an acceptable format (by product, operator, time, dept etc)**
- **Allows team to collect any additional comments during the process**



Data Collection Sheets – why?

- **Consider all the characteristics you want to collect (different inputs and outputs)**
- **Talk to people who are collecting the data (if not you) to explain the reason for this**
- **Test the form before you implement**



Data Collection Sheets – example continuous data

date	operator	product	machine	Size off MC	comments
10/3/21	DJ	ABC123	Press 5	253.25	No issues



Data Collection Sheets – example attribute data

date	operator	product	machine	Scratch Y/N	Coating OK Y/N	comments
10/3/21	DJ	ABC123	Press 5	Y	N	No issues



A well-prepared collection plan will help to ensure successful analysis of the problem by ensuring useful data has been captured

“In God we trust, everybody else bring data”



What is variation?



variation

/vɛːn'eiʃ(ə)n/

See definitions in:

All

Astronomy

Mathematics

Physics

Biology

Music

Ballet

noun

1. a change or slight difference in condition, amount, or level, typically within certain limits.
"regional variations in house prices"

Similar:

difference

dissimilarity

disparity

inequality

contrast

discrepancy



2. a different or distinct form or version of something.
"hurling is an Irish variation of hockey"

Similar:

variant

form

alternative

alternative form

other form

different form



Definitions from Oxford Languages

Feedback

Some terms to help us understand:-

- Mean – commonly called the average
- Median – mid-point of a distribution
- Mode – most frequently occurring number
- Range – difference between maximum value and minimum value
- Standard Deviation – unit of measure of distance from the mean



**Accurate
Precise**



**Not Accurate
Precise**



**Accurate
Not Precise**



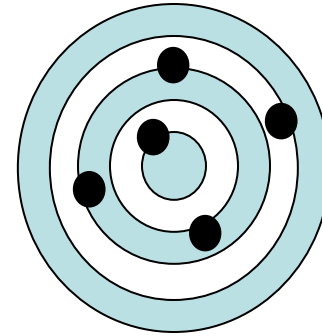
**Not Accurate
Not Precise**

• How's your aim? Are you accurate?

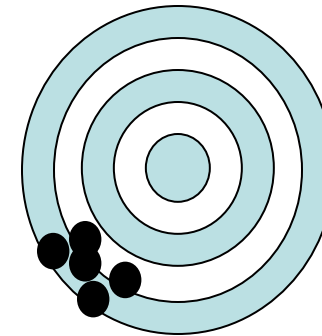
- Is your aim precise; that is, are all the shots clustered around the target, or are they distributed all over the place?

Have you ever...

- Shot a rifle?
- Played darts?
- Shot a round of golf?
- Who is the better shot?



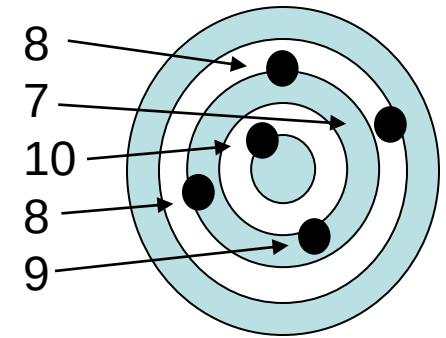
Fred



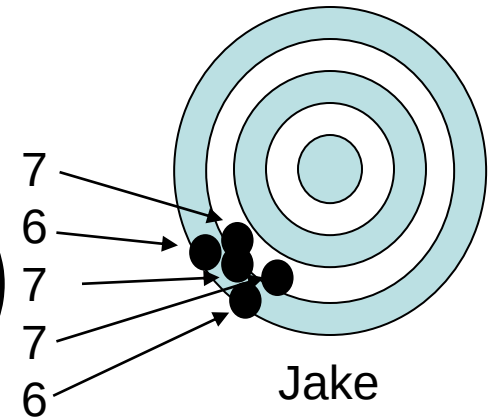
Jake

Deviation

**distance between
observations and
the mean *(or average)***



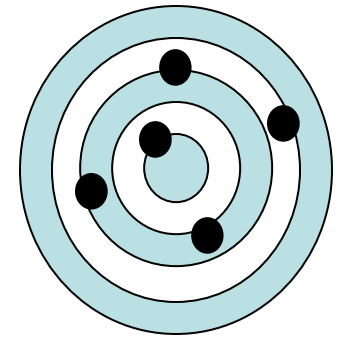
Fred



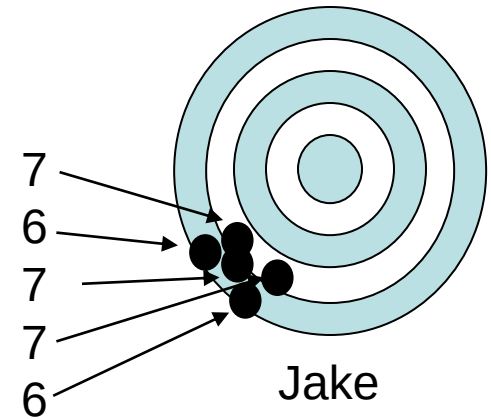
Jake

- Deviation = distance between observations and the mean (or average)

	Observations	Deviations
	7	$7 - 6.6 = 0.4$
	7	$7 - 6.6 = 0.4$
	7	$7 - 6.6 = 0.4$
	6	$6 - 6.6 = -0.6$
	6	$6 - 6.6 = -0.6$
averages	6.6	

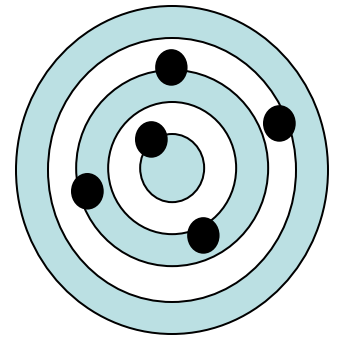


Fred

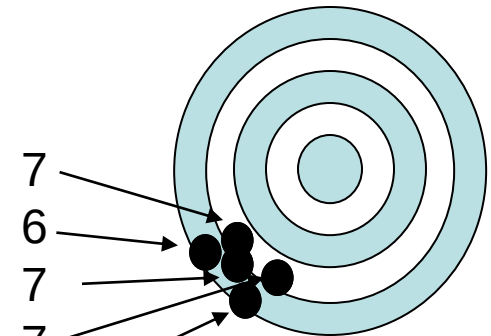


Jake

- Variance = average distance between observations and the mean squared



Fred



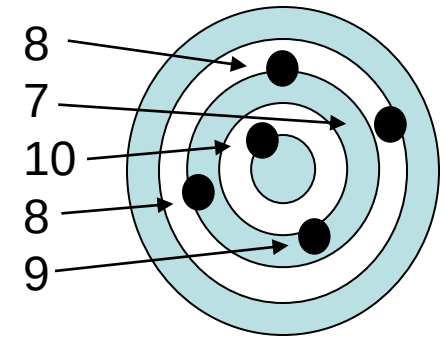
Jake

	Observations	Deviations	Squared Deviations
	7	$7 - 6.6 = 0.4$	0.16
	7	$7 - 6.6 = 0.4$	0.16
	7	$7 - 6.6 = 0.4$	0.16
	6	$6 - 6.6 = 0.6$	0.36
	6	$6 - 6.6 = -0.6$	0.36
averages	6.6		0.24

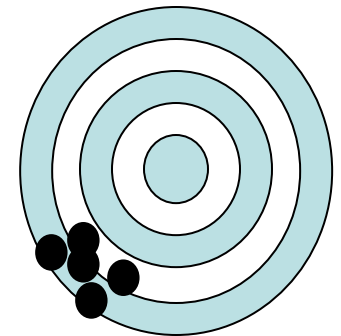
7
6
7
7
6

← **Variance**

- Variance = average distance between observations and the mean squared



Fred



Jake

	Observations	Deviations	Squared Deviations
	10	$10 - 8.4 = 1.6$	2.56
	9	$9 - 8.4 = 0.6$	0.36
	8	$8 - 8.4 = -0.4$	0.16
	8	$8 - 8.4 = -0.4$	0.16
	7	$7 - 8.4 = -1.4$	1.96
averages	8.4		1.0

← Variance

- Standard Deviation is a measure of how close the numbers are to the mean. If the standard deviation is big, then the data is more "dispersed" or "diverse" as it serves as a measure of uncertainty.
- We use Standard Deviation of assessments grades; it gives us an indication of the spread of marks over a module.
- Its symbol is σ (the greek letter sigma)
- The formula is straightforward; it is the **square root** of the **Variance**.

Measure - Standard Deviation

Consider a sample of 5 data points; 6.5, 7.5, 8.0, 7.2, 6.8

The **Range** is the highest minus the lowest; $8.0 - 6.5 = 1.5$

The **Mean** (X-bar) is 7.2

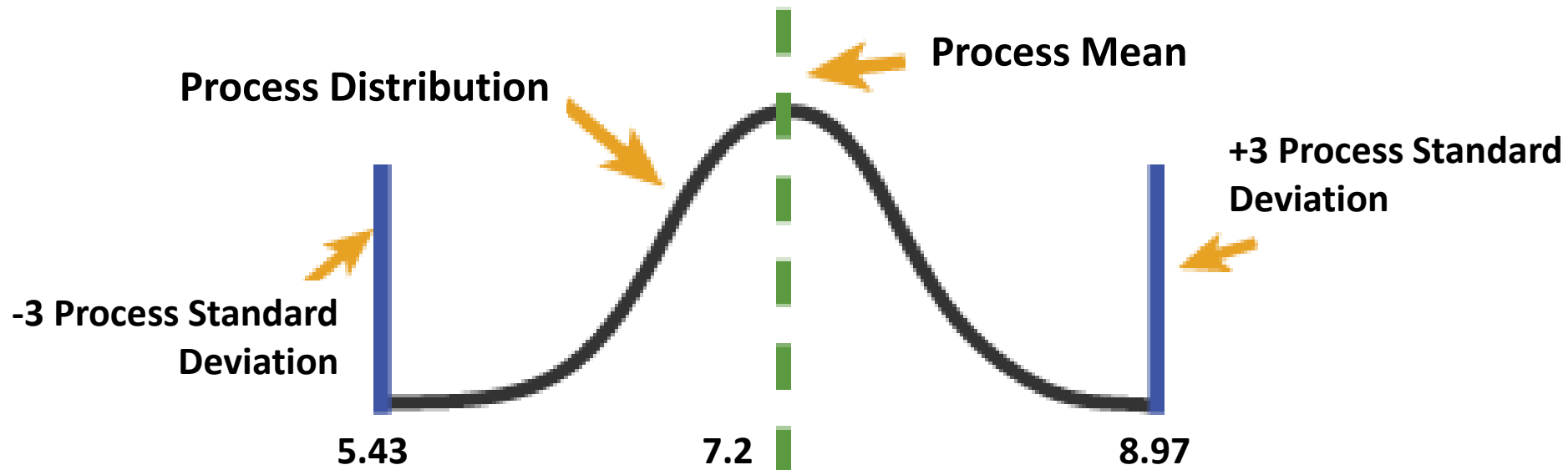
The Standard Deviation is

$$stdev = \sqrt{\frac{(6.5 - 7.2)^2 + (7.5 - 7.2)^2 + (8.0 - 7.2)^2 + (7.2 - 7.2)^2 + (6.8 - 7.2)^2}{5 - 1}}$$

$$stdev = 0.59$$

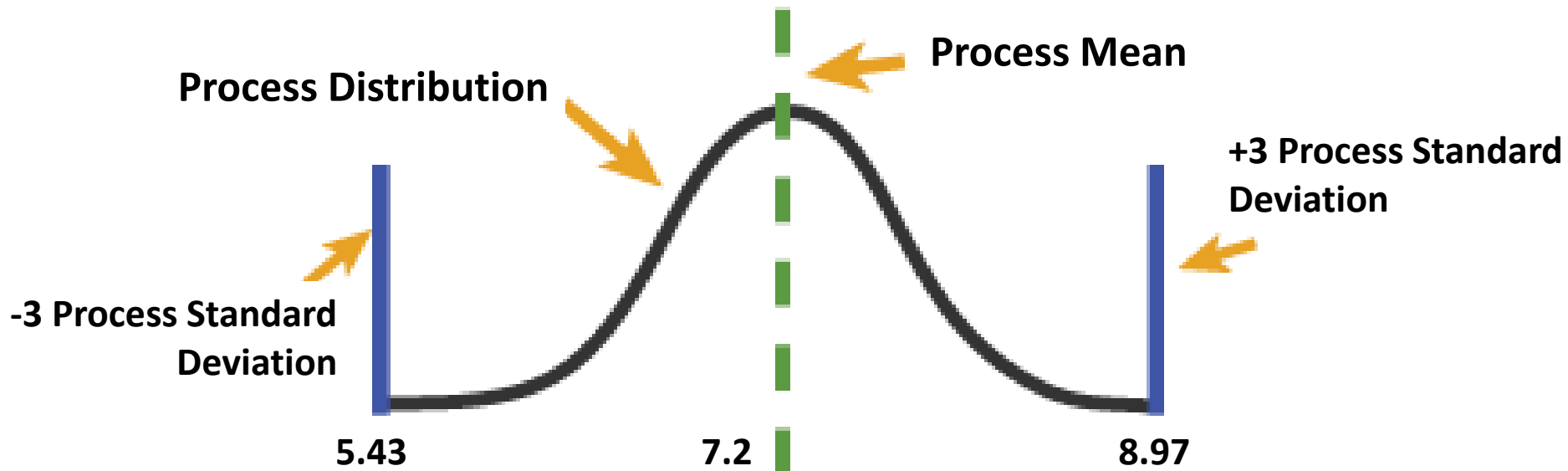
Fortunately, we have Excel!

- Statistical tables have been developed for various types of distributions that quantify the area under the curve for a given number of standard deviations from the mean.
- Shewhart found that control limits placed at three standard deviations from the mean in either direction provide an economical trade-off between the risk of reacting to a false signal and the risk of not reacting to a true signal.

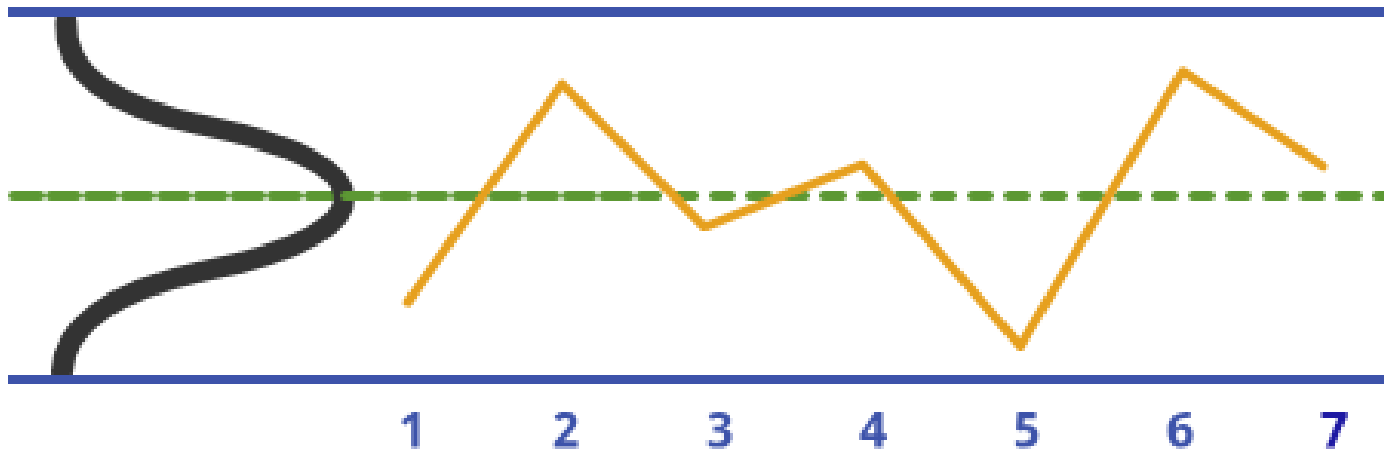


- We want to calculate three-sigma, which is three standard deviations above the mean.
- $3 \times \text{STDEV} + \text{Mean}$
- ***stdev* = 0.59**
→ $(3 \times 0.59) + 7.2 = 8.97$

- The illustration below shows a normal curve for a distribution with a mean of 7.2, a mean less 3 standard deviations value of 5.43, and a mean plus 3 standard deviations value of 8.97. Values, or measurements, less than 5.43 or greater than 8.97 are extremely unlikely.
- These laws of probability are the foundation of the control chart.



- Consider that the distribution is turned sideways, and the lines denoting the mean and ± 3 standard deviations are extended.
- This construction forms the basis of the Control chart.
- Time series data plotted on this chart can be compared to the lines, which now become control limits for the process.
- Comparing the plot points to the control limits allows a simple probability assessment.



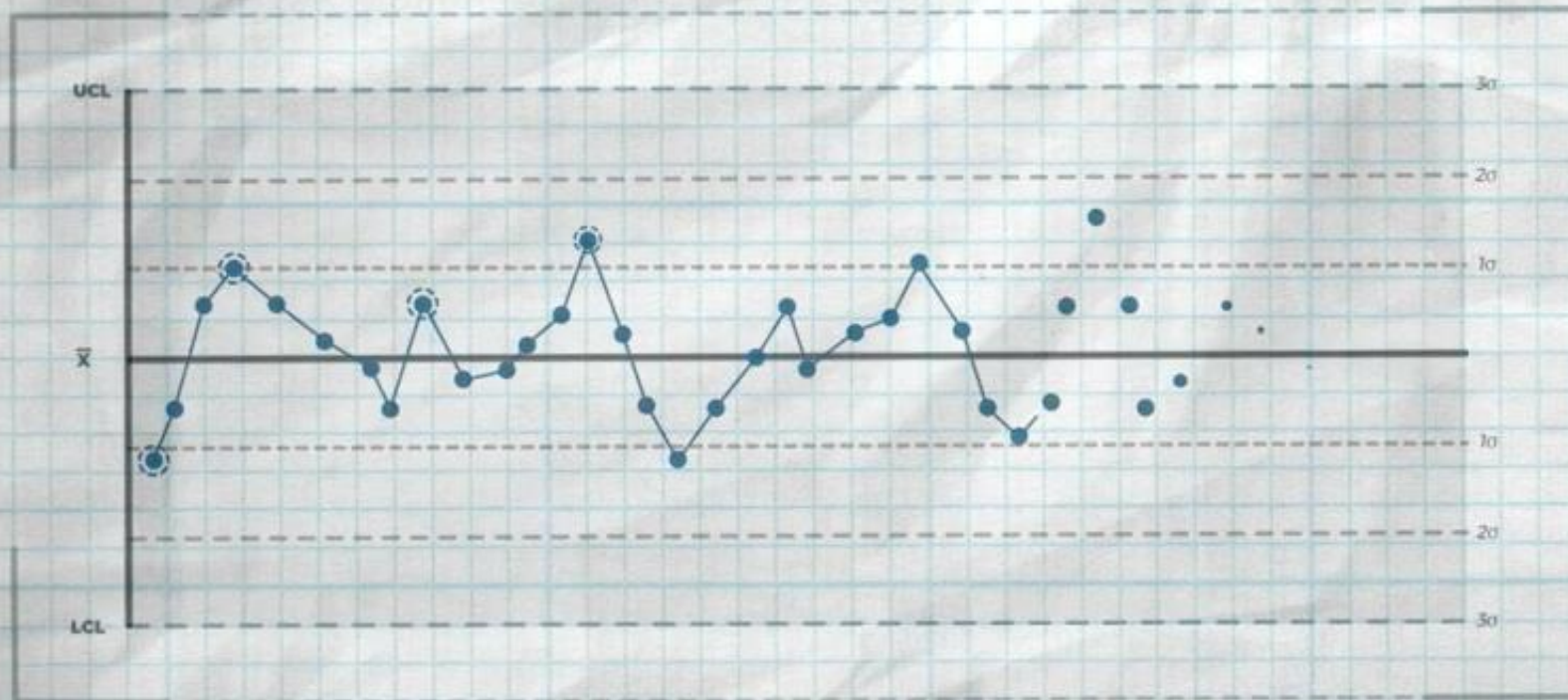
Brief history of (industrial) statistics

Pre 1920	Population & government statistics
1920s	Gossett: t-test (Brewing industry) Fisher: Experimental design (Agriculture)
1930s	W.A. Shewhart: Quality control (Bell Labs) Tippet: application of Fisher's ideas in industry
1940s	Developments in acceptance sampling
1950s	Developments in design of experiments (e.g. George Box) and reliability methods: (e.g. Waloddi Weibull) W. Edwards Deming & Joseph Juran visit Japan
1960s & 1970s	Growth of Japanese industry and their use of statistical techniques including experimental design in quality management. (especially Genichi Taguchi).
1980s & 1990s	‘Quality revolution’ in U.S.A. and Europe. Recognition of the value of statistical techniques and quality management techniques for quality improvement. Many ‘Quality gurus’ established. E.g. W. Edwards Deming, Joseph Juran, Feigenbaum, Crosby, Joiner, Kaoru Ishikawa, Ohno. Many quality programmes and standards (and acronyms!) established: E.g. TQM, JIT, ISO9000, Zero-defects, Six Sigma.

- Control charts are also known as Shewhart charts, named after Walter A. Shewhart, an American physicist, engineer and statistician (1891–1967)
- Control charts were devised as simple, but powerful, graphical ways of monitoring a process to check whether any assignable causes of variation are present.
- Measurements of product quality are made at regular intervals of time and plotted on the vertical axis of the graph versus time on the horizontal axis.
- Prior to this the natural variation of the product will have been estimated and the 'expected' range of natural variation is also shown on the chart - usually as two horizontal lines - called control lines or action lines.

- Points outside the range of natural variation indicate the likely presence of assignable cause variation and appropriate action should then be taken to bring the process back into control.
- Control charts are tools used to determine whether a process is in a state of statistical control or whether special causes of variation are present.
- When there is evidence that special causes of variation are present, action is initiated to bring the process back into control.
- Sometime SPC is used retrospectively to identify where problems occurred and the reasons for these are traced so that these problems may be avoided in the future.

- In SPC, when we talk about the *distribution of the process*, we are referring not to the process itself but to data collected **from** the process—for example, data on widths of pieces coming out of a woodworking machine—and to the way those data are *distributed* when plotted on a chart or graph.



NELSON'S RULES

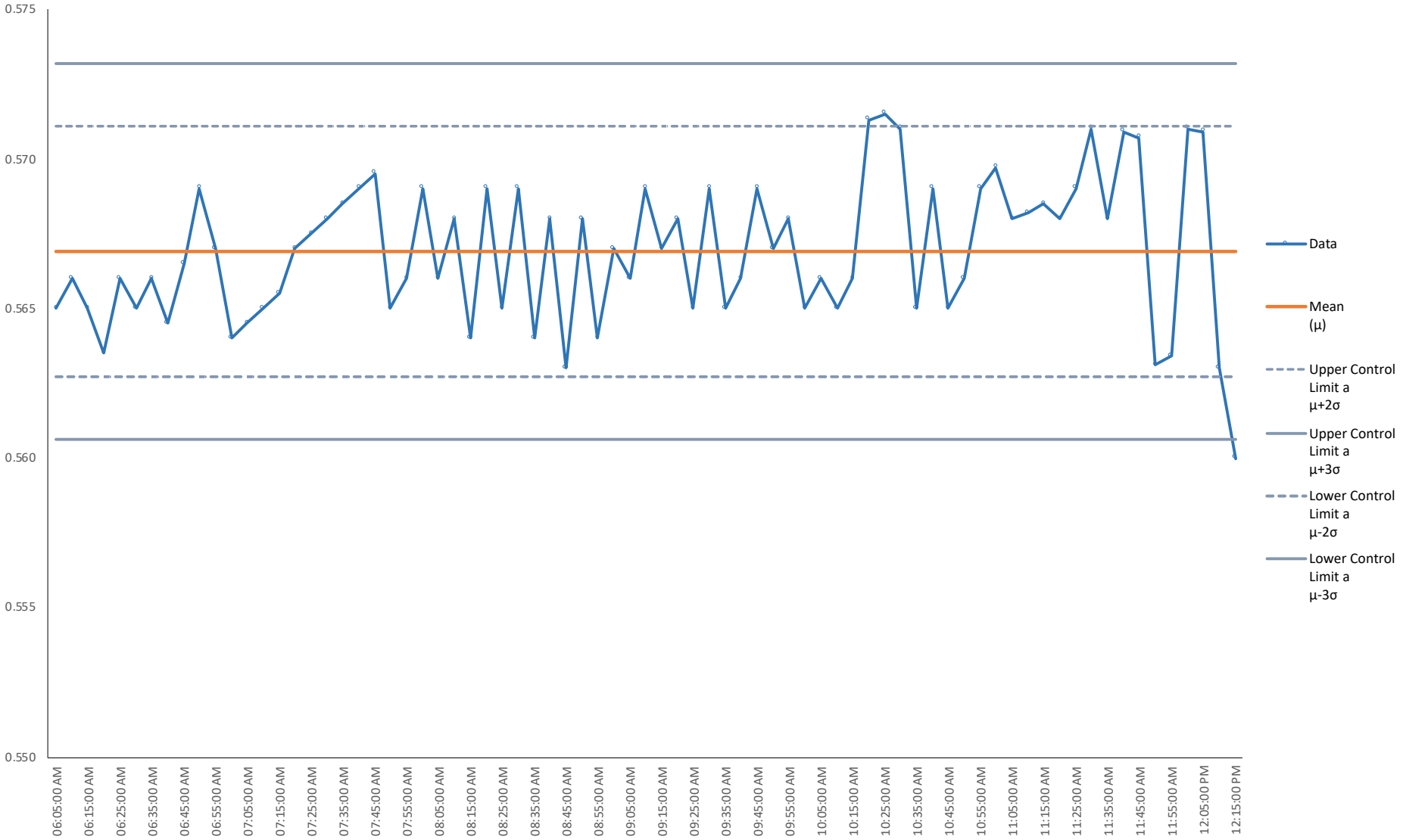
Western Electric Rules

1. One point above UCL or below LCL
2. Two points above/below 2 sigma
3. Four out of five points above/below 1 sigma
4. Eight points in a row above/below the center line

Work of Lloyd S Nelson, published in October 1984;

1. One point above UCL or below LCL
2. Two points above/below 2 sigma
3. Three out of four points above/below 1 sigma
4. Eight points in a row above/below the center line
5. Six points in a row ascending or descending (trend)
6. 15 points in a row "hugging" the center line (between -1 and +1 sigma)
7. 14 points in a row alternating up and down
8. Eight points in a row above 1 sigma or below -1 sigma

Control Chart



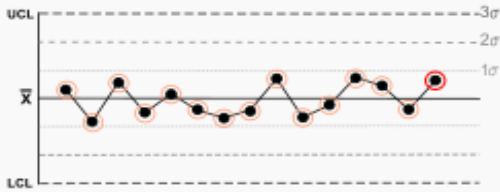
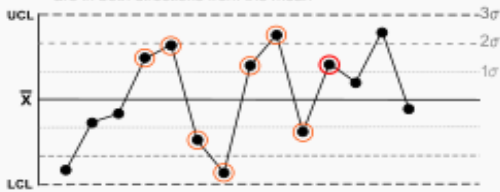
Nelson's Rules 1 to 3

Rule	Description	Chart Example	Problem Indicated
Rule 1	One point is more than 3 standard deviations from the mean.	Rule 1: One point is more than 3 standard deviations from the mean	One sample (two shown in this case) is grossly out of control.
Rule 2	Nine (or more) points in a row are on the same side of the mean.	Rule 2: Nine (or more) points in a row are on the same side of the mean	Some prolonged bias exists.
Rule 3	Six (or more) points in a row are continually increasing (or decreasing).	Rule 3: Six (or more) points in a row are continually increasing (or decreasing)	A trend exists.

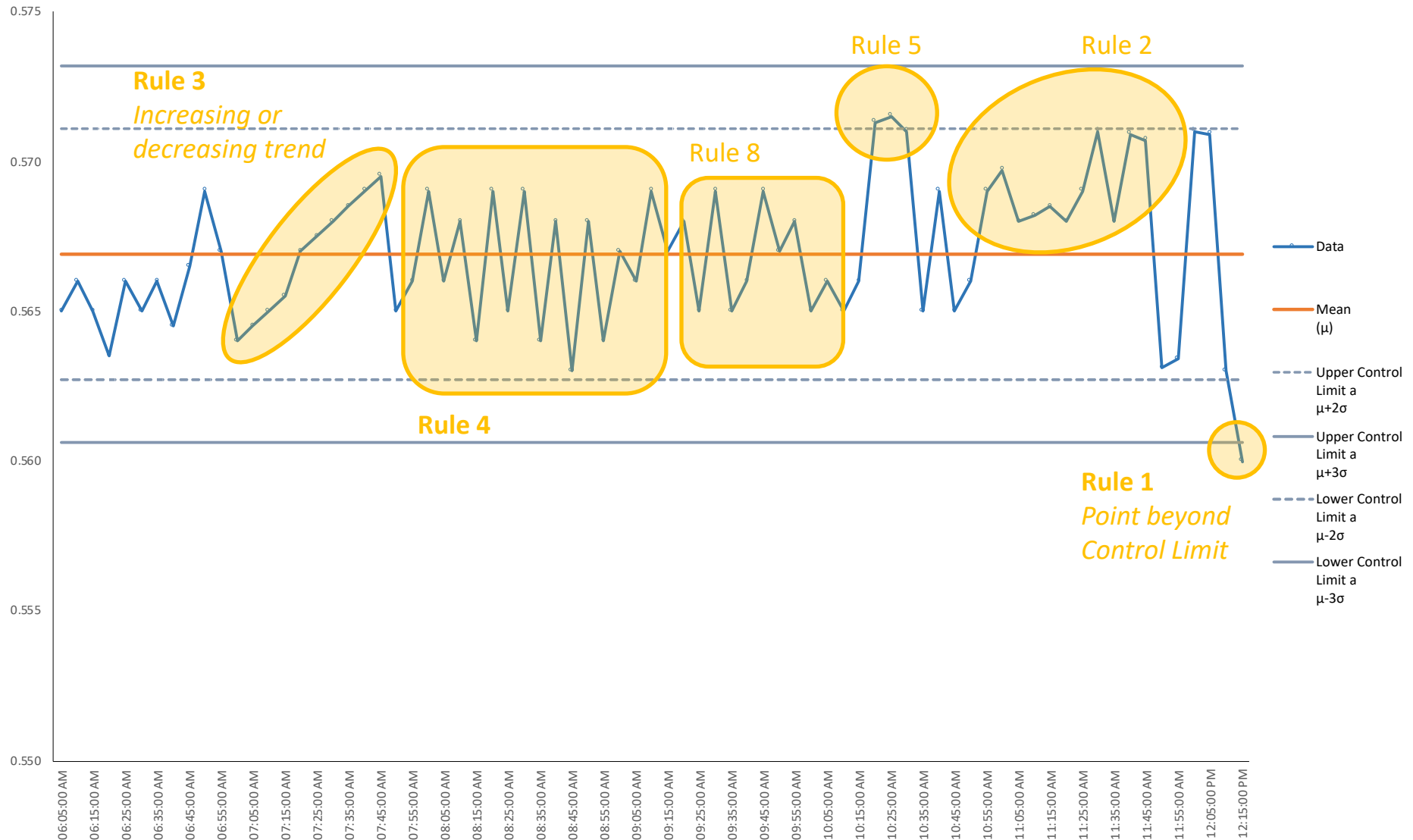
Nelson's Rules 4 to 6

Rule 4	Fourteen (or more) points in a row alternate in direction, increasing then decreasing.	Rule 4: Fourteen (or more) points in a row alternate in direction, increasing then decreasing	This much oscillation is beyond noise. This is directional and the position of the mean and size of the standard deviation do not affect this rule.
Rule 5	Two (or three) out of three points in a row are more than 2 standard deviations from the mean in the same direction.	Rule 5: Two (or three) out of three points in a row are more than 2 standard deviations from the mean in the same direction	There is a medium tendency for samples to be mediumly out of control. The side of the mean for the third point is unspecified.
Rule 6	Four (or five) out of five points in a row are more than 1 standard deviation from the mean in the same direction.	Rule 6: Four (or five) out of five points in a row are more than 1 standard deviation from the mean in the same direction	There is a strong tendency for samples to be slightly out of control. The side of the mean for the fifth point is unspecified.

Nelson's Rules 7 and 8

Rule 7	Fifteen points in a row are all within 1 standard deviation of the mean on either side of the mean.	Rule 7: Fifteen points in a row are all within 1 standard deviation of the mean on either side of the mean 	With 1 standard deviation, greater variation would be expected.
Rule 8	Eight points in a row exist with none within 1 standard deviation of the mean and the points are in both directions from the mean.	Rule 8: Eight points in a row exist with none within 1 standard deviation of the mean and the points are in both directions from the mean 	Jumping from above to below whilst missing the first standard deviation band is rarely random.

Control Chart



Process in Control when...

- No sample points outside the limits
- Most points close to process average.
- An equal number of points above and below the centerline.
- All points appear randomly distributed.

MADE

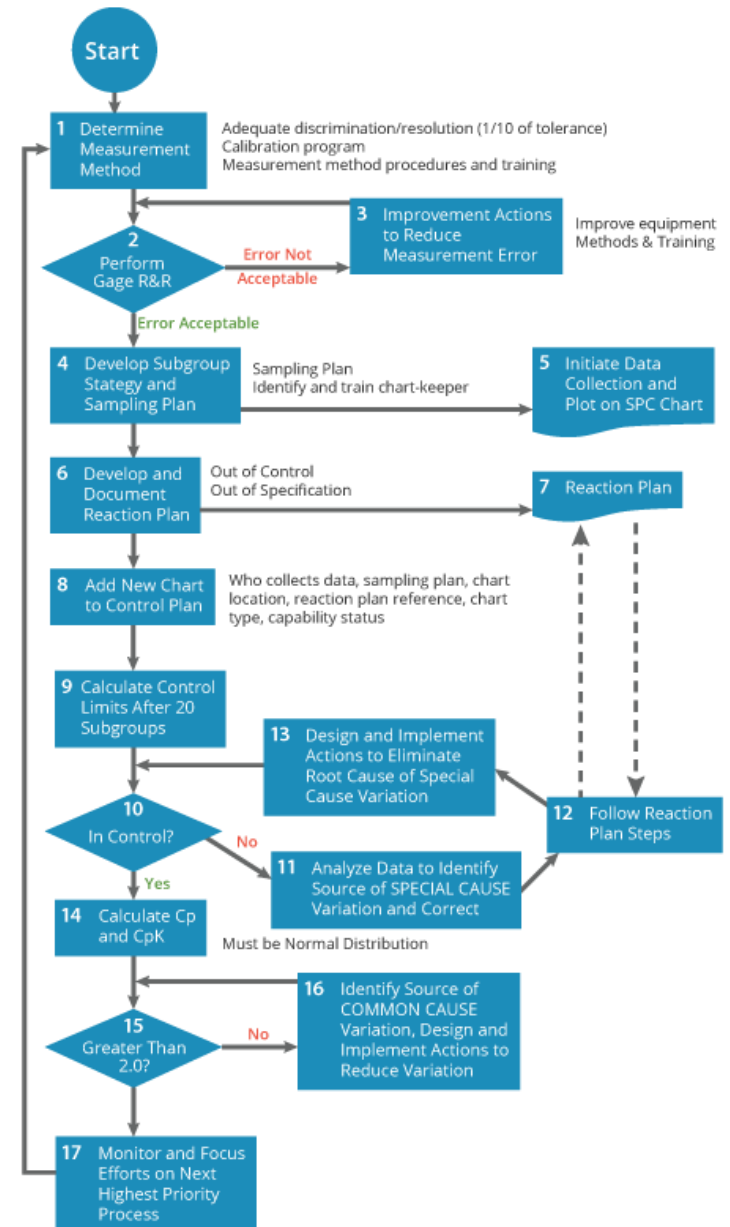
Time Period	Data	Mean (d)	Moving range	Average of moving ranges	Upper Control Limit (u+3s)	Upper Control Limit (u+2s)	Lower Control Limit (u-2s)	Lower Control Limit (u-3s)
6:05:00 AM	0.563	0.567	0.000	0.000	0.571	0.571	0.565	0.560
6:10:00 AM	0.566	0.567	0.001	0.000	0.571	0.571	0.565	0.560
6:15:00 AM	0.567	0.567	0.001	0.000	0.571	0.571	0.565	0.560
6:20:00 AM	0.564	0.567	0.001	0.000	0.571	0.571	0.565	0.560
6:25:00 AM	0.566	0.567	0.002	0.000	0.571	0.571	0.565	0.560
6:30:00 AM	0.565	0.567	0.001	0.000	0.571	0.571	0.565	0.560
6:35:00 AM	0.566	0.567	0.001	0.000	0.571	0.571	0.565	0.560
6:40:00 AM	0.565	0.567	0.001	0.000	0.571	0.571	0.565	0.560
6:45:00 AM	0.567	0.567	0.002	0.000	0.571	0.571	0.565	0.560
6:50:00 AM	0.568	0.567	0.002	0.000	0.571	0.571	0.565	0.560
6:55:00 AM	0.567	0.567	0.002	0.000	0.571	0.571	0.565	0.560
7:00:00 AM	0.564	0.567	0.002	0.000	0.571	0.571	0.565	0.560
7:05:00 AM	0.565	0.567	0.001	0.000	0.571	0.571	0.565	0.560
7:10:00 AM	0.565	0.567	0.000	0.000	0.571	0.571	0.565	0.560
7:15:00 AM	0.566	0.567	0.001	0.000	0.571	0.571	0.565	0.560
7:20:00 AM	0.567	0.567	0.001	0.000	0.571	0.571	0.565	0.560
7:25:00 AM	0.568	0.567	0.001	0.000	0.571	0.571	0.565	0.560
7:30:00 AM	0.568	0.567	0.000	0.000	0.571	0.571	0.565	0.560
7:35:00 AM	0.569	0.567	0.001	0.000	0.571	0.571	0.565	0.560
7:40:00 AM	0.569	0.567	0.000	0.000	0.571	0.571	0.565	0.560
7:45:00 AM	0.570	0.567	0.001	0.000	0.571	0.571	0.565	0.560
7:50:00 AM	0.565	0.567	0.005	0.000	0.571	0.571	0.565	0.560
7:55:00 AM	0.566	0.567	0.001	0.000	0.571	0.571	0.565	0.560
8:00:00 AM	0.569	0.567	0.002	0.000	0.571	0.571	0.565	0.560
8:05:00 AM	0.566	0.567	0.003	0.000	0.571	0.571	0.565	0.560
8:10:00 AM	0.568	0.567	0.002	0.000	0.571	0.571	0.565	0.560
8:15:00 AM	0.564	0.567	0.004	0.000	0.571	0.571	0.565	0.560
8:20:00 AM	0.569	0.567	0.005	0.000	0.571	0.571	0.565	0.560
8:25:00 AM	0.565	0.567	0.006	0.000	0.571	0.571	0.565	0.560
8:30:00 AM	0.569	0.567	0.004	0.000	0.571	0.571	0.565	0.560
8:35:00 AM	0.564	0.567	0.005	0.000	0.571	0.571	0.565	0.560
8:40:00 AM	0.568	0.567	0.006	0.000	0.571	0.571	0.565	0.560
8:45:00 AM	0.563	0.567	0.003	0.000	0.571	0.571	0.565	0.560
8:50:00 AM	0.568	0.567	0.005	0.000	0.571	0.571	0.565	0.560
8:55:00 AM	0.564	0.567	0.004	0.000	0.571	0.571	0.565	0.560
9:00:00 AM	0.567	0.567	0.000	0.000	0.571	0.571	0.565	0.560
9:05:00 AM	0.566	0.567	0.001	0.000	0.571	0.571	0.565	0.560
9:10:00 AM	0.569	0.567	0.005	0.000	0.571	0.571	0.565	0.560
9:15:00 AM	0.567	0.567	0.002	0.000	0.571	0.571	0.565	0.560
9:20:00 AM	0.568	0.567	0.001	0.000	0.571	0.571	0.565	0.560
9:25:00 AM	0.565	0.567	0.005	0.000	0.571	0.571	0.565	0.560





Implementing Statistical Process Control

- Deploying Statistical Process Control is a process in itself, requiring organizational commitment across functional boundaries.
- The flow-chart below outlines the major components of an effective SPC effort.



- Different types of control charts depending on types of data and sample sizes

Variable (measureable, continuous data)

- Mean/Range - ($\bar{x}$ - and R- chart) – sample size up to 10
- Individual/Range (ImR – chart) – best with sample size of one
- Mean/Std Dev – ($\bar{x}$ and S) – larger sub-group sizes over 10

Attribute data

- P-chart/NP-chart/C-chart/U-chart

Variable (measureable, continuous data)

- Plots the average or individual points and the range
- Mean/Std Dev – ($\bar{x}$ and S) – larger sub-group sizes over 10

Attribute data

- Plots the proportion or count rather than variable data
- P chart – charts the proportion of defects
- C chart – charts defect count per sample (same sample size each time)
- U chart – charts number of defects per unit (sample size can vary)
- NP chart – number of defects in each sub-group

Lots of other charts exist

P chart	C charts	U charts	Cusum & EWMA
For yes-no questions like “is it defective?” (binomial data)	For counting number defects where most items have ≥ 1 defects (eg. custom built houses)	Average count per unit (similar to C chart)	Advanced charts
$\bar{p} \pm 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$	$\bar{c} \pm 3\sqrt{\bar{c}}$	$\bar{u} \pm 3\sqrt{\frac{\bar{u}}{n}}$	“V” shaped or Curved control limits (calculate them by hiring a statistician)

- Chosen so that variation within the sample is considered to be from common causes
- Special causes should only occur between samples
- Special causes to avoid in sampling
 - passage of time
 - workers
 - shifts
 - machines
 - Locations

- Larger samples are more accurate
- Sample costs money, but so does being out-of-control
- Don't convert measurement data to "yes/no" binomial data (X's to P's)
- Don't combine data (or mix product)
- Have out-of-control procedures (what do I do now?)
- Actual production volume matters (Average Run Length)

- Nature of defect is different in services.
- Service defect is a failure to meet customer requirements.
- Monitor times, customer satisfaction.

Hospitals?

- Timeliness and quickness of care, staff responses to requests, accuracy of lab tests, cleanliness, courtesy, accuracy of paperwork, speed of admittance.

Airlines?

- Flight delays, lost luggage and luggage handling, waiting time at ticket counters and check-in, agent and flight attendant courtesy, accurate flight information, passenger cabin cleanliness.

Level 5 – Continuous Improvement with Industry 4.0

Measure 3 – Capability





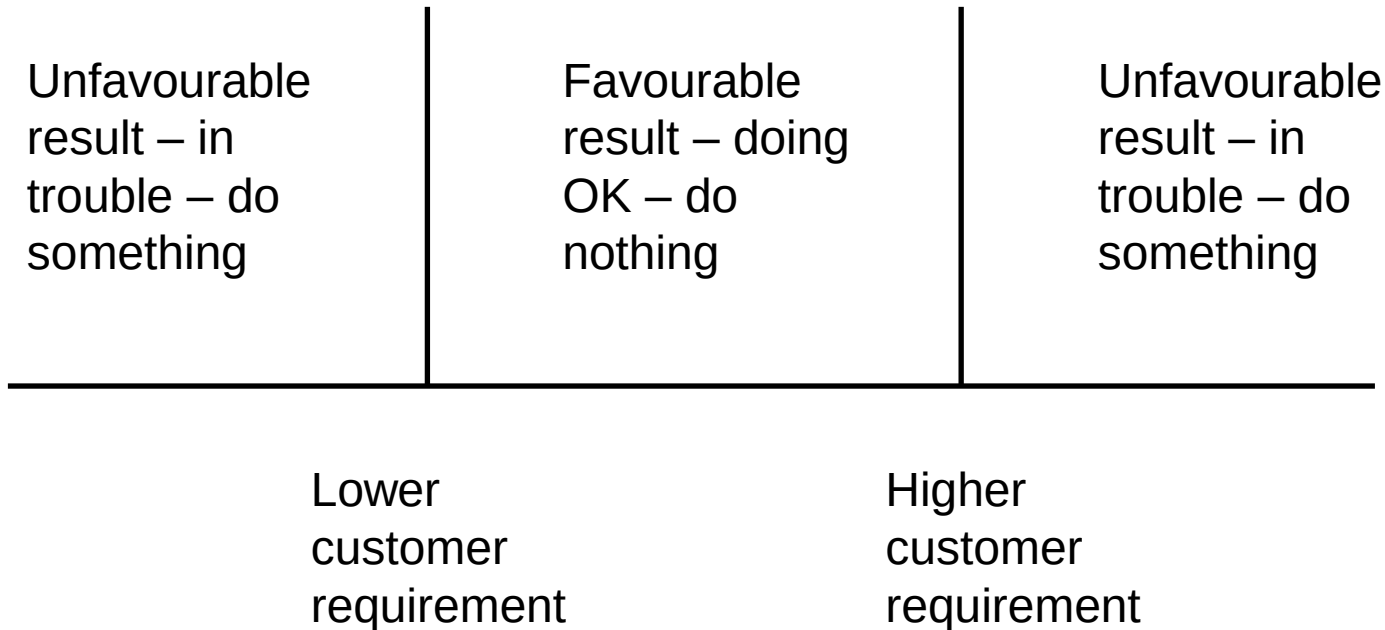
There are two methods for interpreting data within a manufacturing environment:-

The Specification Value approach (VoC)

Shewharts Control Chart approach (Voice of the Process)



The Specification Value approach (VoC)



- In manufacturing, the questions are:
 - Where is the process centered; in other words, is the process on or off target and, if it's the latter, by how much is it off target?
 - How much does the process fluctuate about the center?

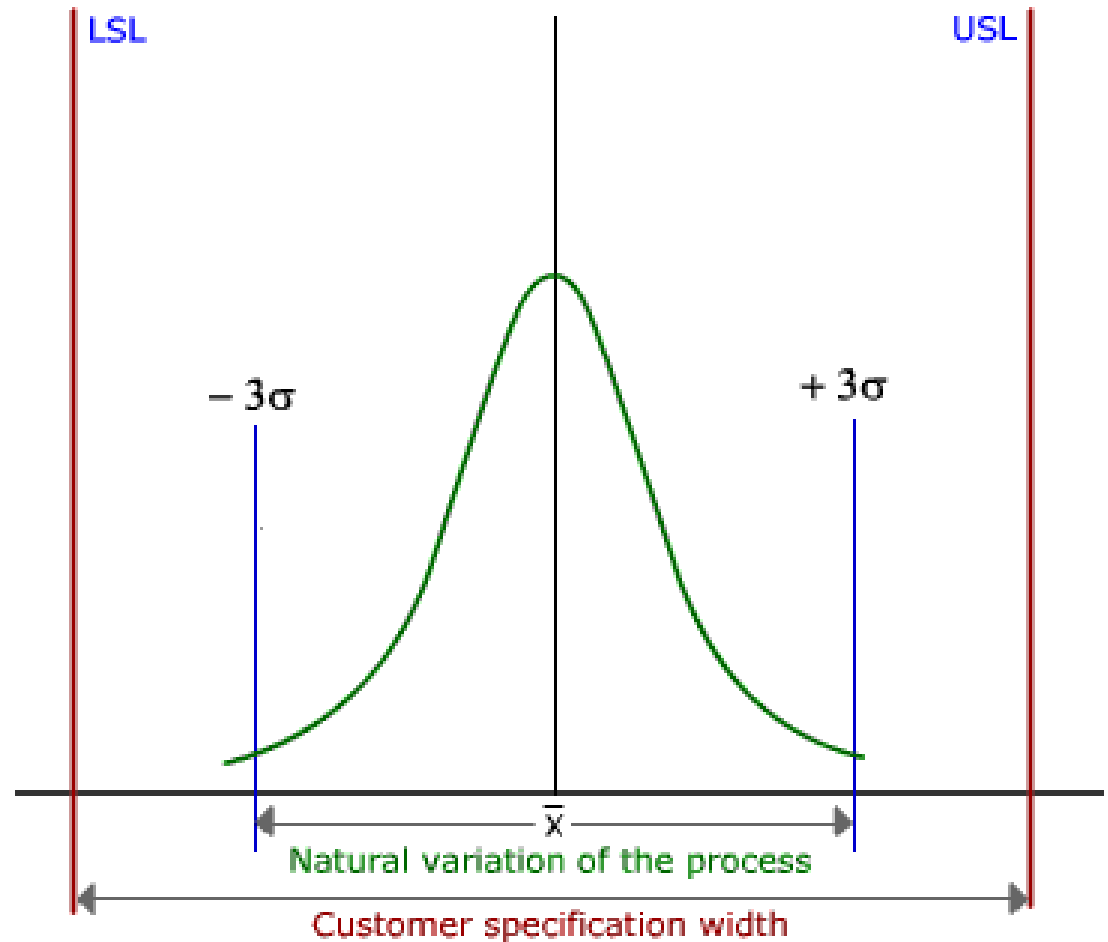
- Process and Control limits (VoP):
 - Statistical
 - Process limits are used for individual items
 - Control limits are used with averages
 - Limits = $\mu \pm 3\sigma$
 - Define usual (common causes) & unusual (special causes)
- Specification limits (VoC):
 - Engineered
 - Limits = target $\pm$ tolerance
 - Define acceptable & unacceptable

Control limits, are the 'Voice of the Process' Identified as;

- Upper and Lower Control Limits

Tolerances are part of the Customer specification, and therefore the 'Voice of the Customer' Identified as;

- Upper and Lower Specification Limits (LSL & USL)

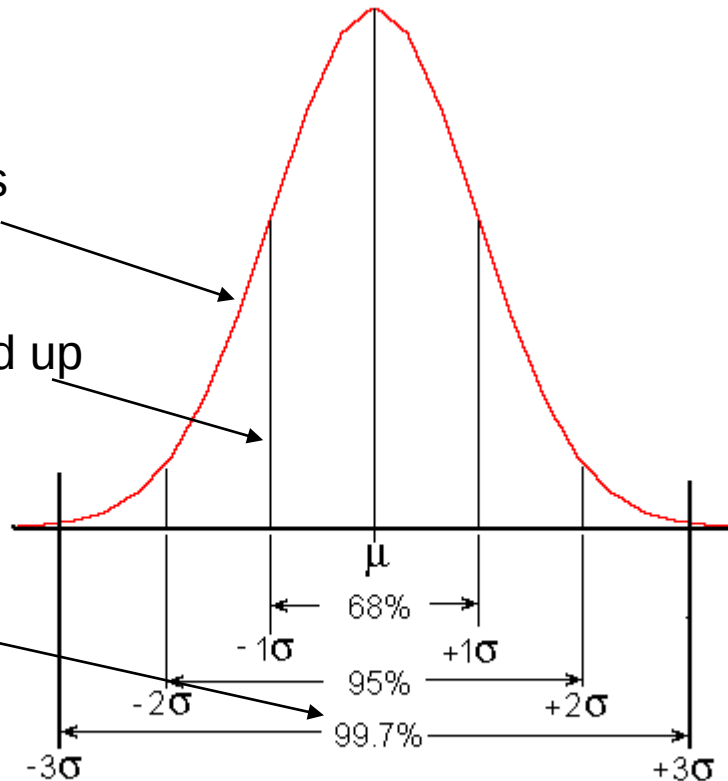


“Normal” bell shaped curve

Add up about 30 of most things
and you start to be “normal”

Normal distributions are divided up
into 3 standard deviations on
each side of the mean

Once you have that, you
know a lot about
what is going on

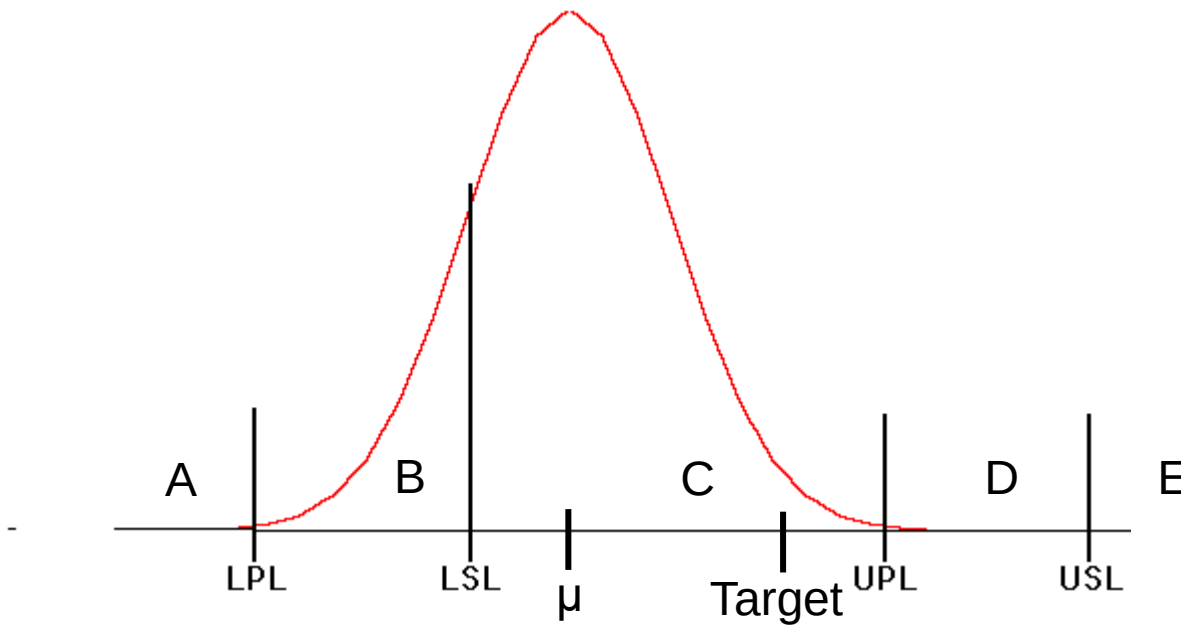


And that is what a standard deviation
is good for



- Common Causes:
 - Random variation (usual)
 - No pattern
 - Inherent in process
 - adjusting the process increases its variation
- Special Causes
 - Non-random variation (unusual)
 - May exhibit a pattern
 - Assignable, explainable, controllable
 - adjusting the process decreases its variation

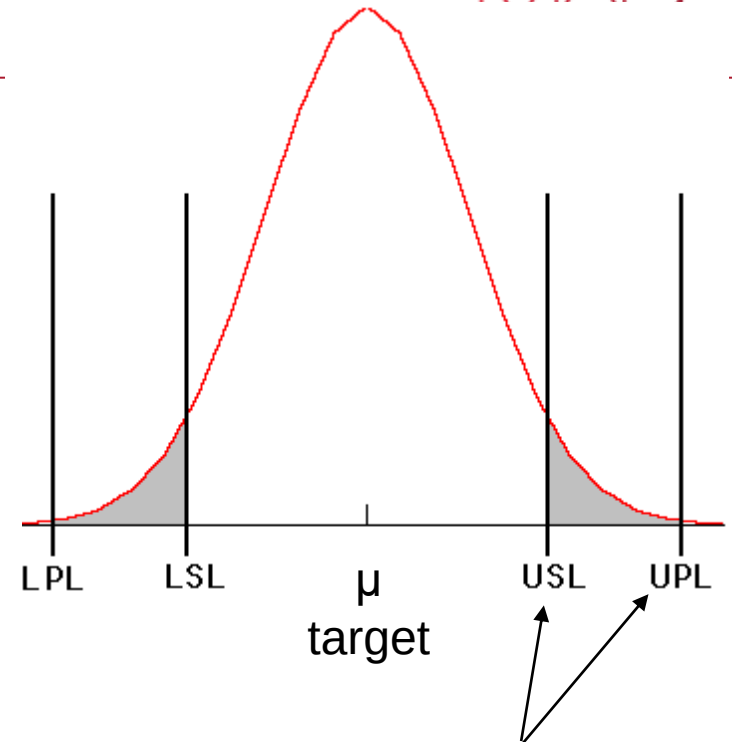
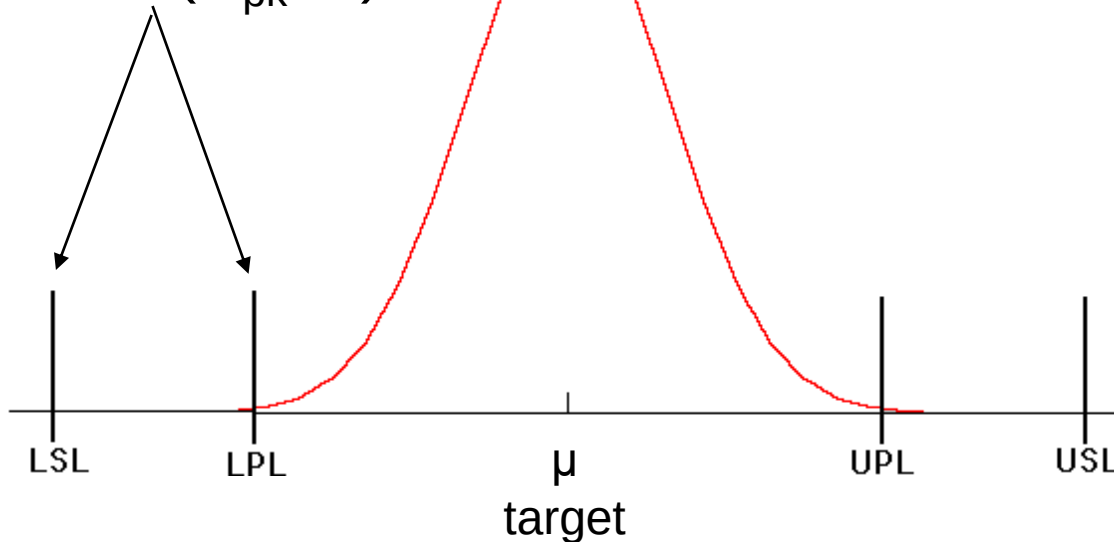
Usual v. Unusual, Acceptable v. Defective



More about limits



Good quality:
defects are
rare ($C_{pk} > 1$)



Poor quality:
defects are
common ($C_{pk} < 1$)

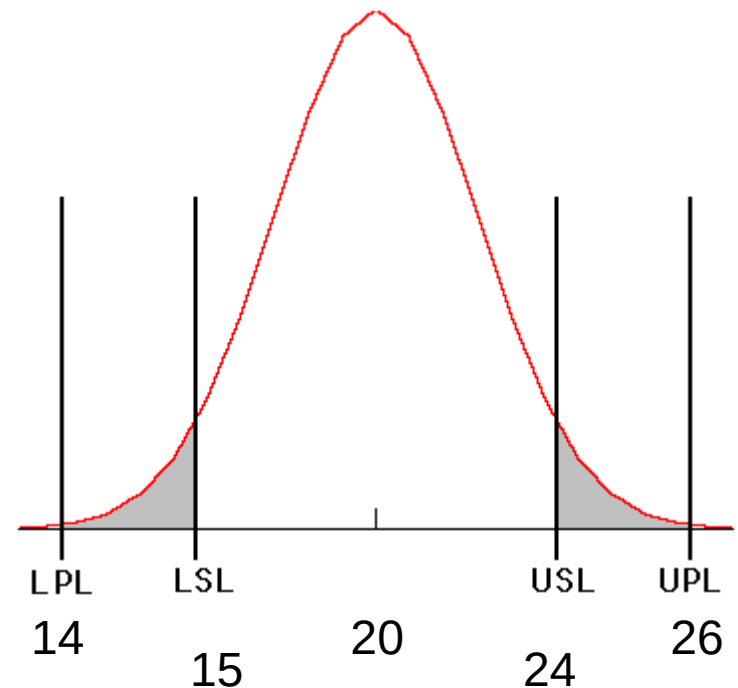
C_{pk} measures “Process Capability”

If process limits and control limits are at the same location, $C_{pk} = 1$. $C_{pk} \geq 2$ is exceptional.

Good quality: defects are rare ($C_{pk} > 1$)

Poor quality: defects are common ($C_{pk} < 1$)

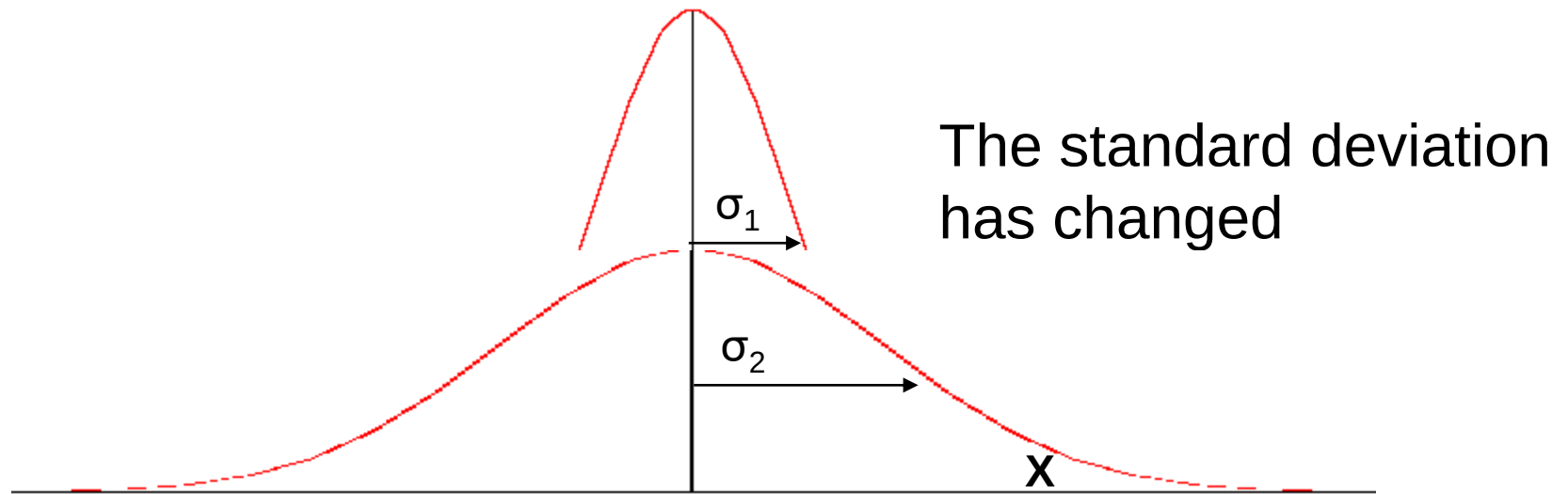
$$C_{pk} = \min \left\{ \begin{array}{l} \frac{USL - \bar{\bar{x}}}{3\sigma} = \frac{24 - 20}{6} = .667 \\ \frac{\bar{\bar{x}} - LSL}{3\sigma} = \frac{20 - 15}{6} = .833 \end{array} \right.$$



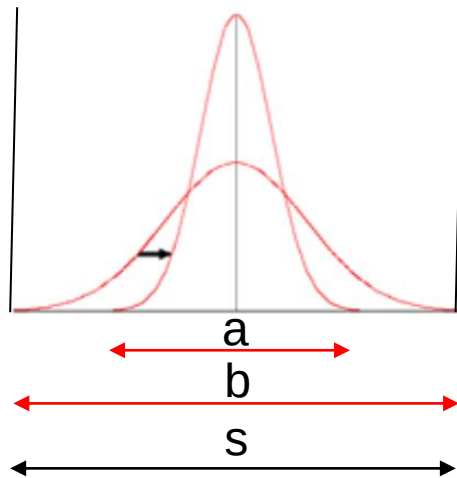
$$3\sigma = (UPL - \bar{\bar{x}}, \text{ or } \bar{\bar{x}} - LPL)$$

- When an observation is unusual, what can we conclude?

- When an observation is unusual, what can we conclude?

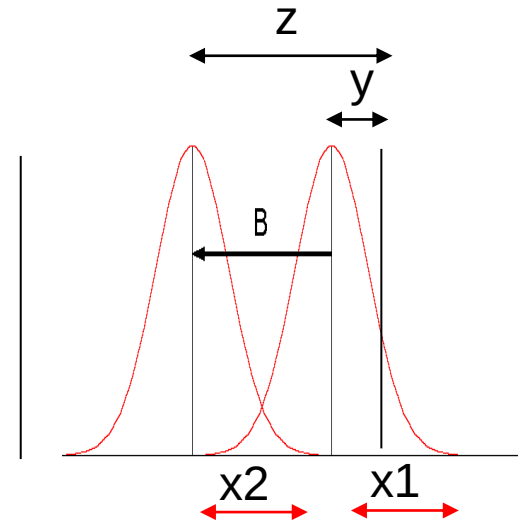


- We can now begin to recognise process drift and changes and give them values



$$\begin{aligned} a/s &= \\ Cp &2 \\ b/s &= \\ Cp &1 \end{aligned}$$

Cp measures process spread in relation to specification spread



$$\begin{aligned} y/x1 &= \\ Cpk &.5 \\ z/x2 &= \\ \text{approx} &3 \end{aligned}$$

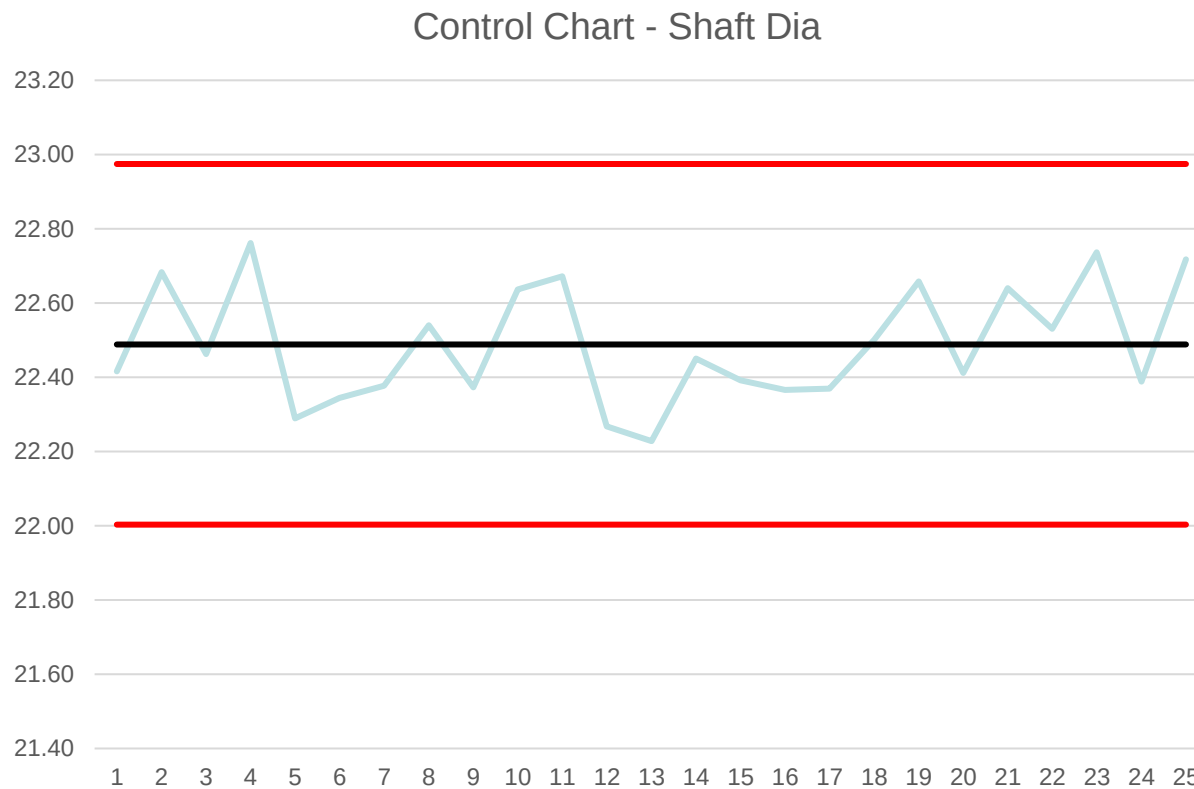
Cpk measures process spread in relation to specification spread and its location relative to process mean and closest spec limit

Exercise

- Open the spreadsheet Measure 3 – capability exercise

measurements	average	diff	sq diff		
22.42			0.000		
22.68			0.000		
22.46			0.000		
22.76			0.000		
22.29			0.000		
22.34			0.000		
22.38			0.000		
22.54			0.000		
22.37			0.000		
22.64			0.000		
22.67			0.000		
22.27			0.000		
22.23			0.000		
22.45			0.000		
22.39			0.000		
22.37			0.000		
22.37			0.000		
22.50			0.000		
22.66			0.000		
22.41			0.000		
22.64			0.000		
22.53			0.000		
22.74			0.000		
22.39			0.000		
22.72			0.000		
ave			0.000	variance	
0.00UCL			0.000	std dev	
0.00LCL					
lower		specification			
22.00spec		tol			
23.00upper spec		1.00	spread	0	#DIV/0! Cp
					#DIV/0! CpU
					#DIV/0! CpL

- Open the spreadsheet Measure 3 – capability exercise



- C_p will give you the process capability against the specification but will not give you its location
- C_{pk} will give you the location to one side of the process in relation to a specification limit but will not give you the other side performance
- You have to use both C_p and C_{pk} to get the full picture!

- 25/11 and 2/12 normal teaching
- 09/12 no teaching – I will be available to review your project and any other information (i.e. SIPOC, VSM)
- 16/12 – feedback to group
 - 5 min overview on your company
 - What your project is about
 - Draft Project Charter