Chapter 19 – Process Capability

If I asked you . . . . . How capable is your process of producing parts that meet your customers needs.

Would you know how to respond?

One question you may ask is. . . .

How do you conclude that a process is "capable" of producing a product?

This is where the concept of **Process Capability Analysis** comes into play.

This chapter is dedicated to this topic and has 8 major sections:

Section 1 - **What is Process Capability Analysis & Why It Matters**

Section 2 - Understanding the **Prerequisites of** a **process capability study**

Section 3 - How to calculate the **4 major process capability indices** including C_p, C_{pk}, C_{pm} & C_r .

Section 4 - Understanding the **2 major process performance indices** including P_p, P_{pk}.

Section 5 - How to **interpret the results** & the **5 Reactions** to your process capability study.

Section 6 – Process Capability for **Attribute Data**

Section 7 – Process Capability for **Non-Normal Data**

Section 8 – Process Performance **Metrics** (% Defective, PPM, DPMO, DPU, FPY, RTY).

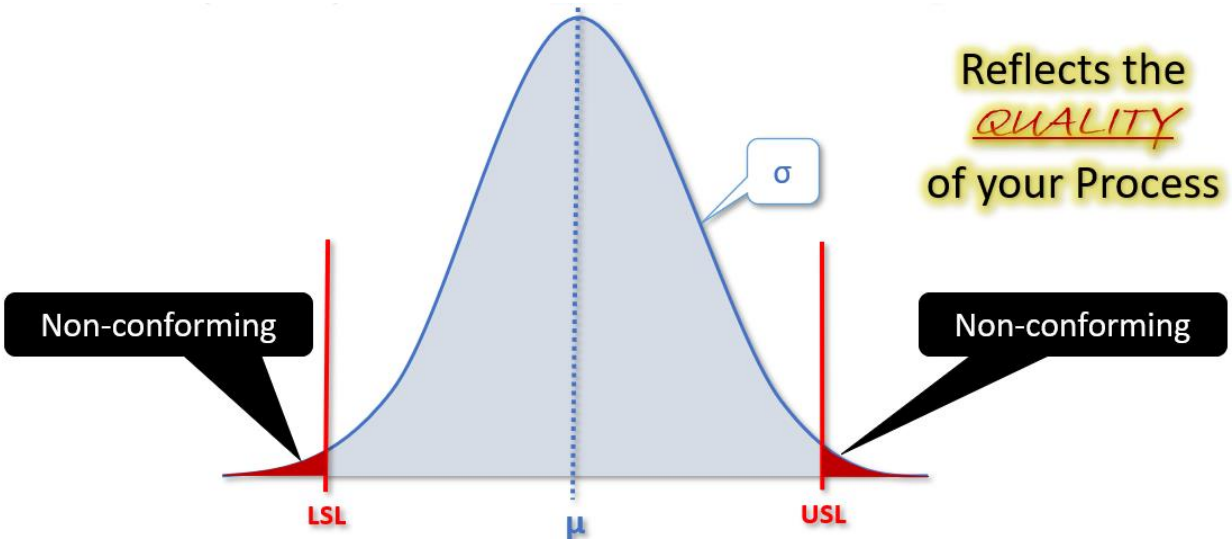
What is Process Capability Analysis & Why It Matters

How often do we have to build good product before a process can be confirmed as being "capable"?

To answer this question, Continuous Improvement Professionals have developed the **Process Capability Analysis Tool**.

Process Capability Analysis *allows us to quantify the capability of our process to* ***produce*** *product that* ***meets the design specifications****.*

More specifically, this tool can be used to predict what portion of the overall population of product produced will fall outside of the customers specification limits - and thus ***result in a defect***.



This ability to make predictions is an extremely powerful one and is the primary benefit of Process Capability Analysis.

Process Capability Analysis can also be used in other instances as well.

For example, *Process Capability Analysis can be the starting point of a continuous improvement project*.

In this way, The Process Capability Analysis can establish the baseline for your process; and as you'll learn below, it can help guide you in how your process needs to be improved.

Process Capability Analysis can also be used on the back end of a project to measure the effect of a change on a process.

We will talk more below about the relationship between process variation and process capability and I'll comment quickly that the idea of six standard deviations (sigma) of capability was the foundation of the **Six-Sigma** concept.

Understanding the Difference Between Process Specifications & Process Performance

Process Capability can be described as a comparison of your **process performance** against its **process specifications** using various *capability indices*.

To better understand these concepts, I've broken this section down into 3 sub-sections.

In the first two sections we will define **Process Performance** & **Process Specifications**.

Then the third section explains their relationship with **Process Capability Analysis,** which at the end of the day looks like this:

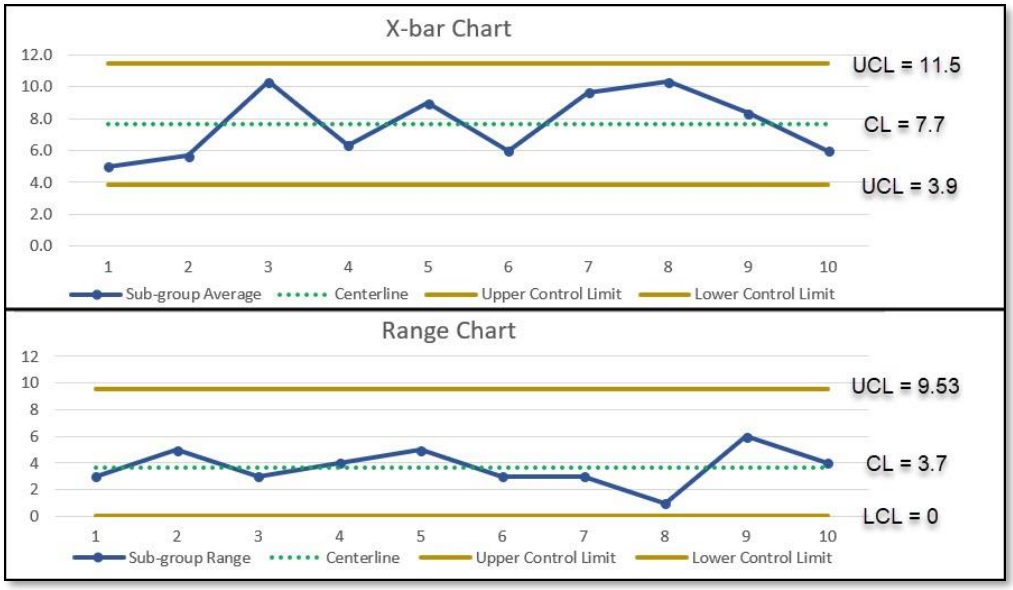
$$\boldsymbol{Process\ Capability\ Analysis = \frac{Process\ Specification}{Process\ Performance}}$$

Process Performance

Process Performance is the natural ***variability*** associated with your **stable** process.

Remember a "stable" process cannot be experiencing any special cause variation; only normal cause variation.

Below I'm showing a control chart from a process that's only experiencing natural variability and thus is "in control".



It should be understood that *it is possible to have a process that is in control and also producing bad product* - which is where process capability analysis comes into play.

The width of this process is considered the process performance. This process width reflects the natural variation associated with the process.

A process with less variability could be described as having have higher performance and a low performance process has high variability.

Process Specifications

Process Specifications are the ***pre-defined specifications*** associated with the product that's produced by your process.

For example, if you're building a widget (example below) that has critical features (highlighted), you would expect to find that critical feature on the product drawing with a dimension and tolerance.

That dimension & tolerance can be considered the product specification and is usually determined by R&D or your customer directly:

Because this is a critical feature of the product it generally has some relationship with the functionality of the product and therefore, whenever a product is built outside of those tolerance the product will not function as intended.

Also, this tolerance implies that the design can tolerate that level of variability before functionality is affected.

Below you'll see two examples how why the center of your process matters a lot for process capability.

Within the context of this chapter the Process Specification is generally defined by the **Lower Specification Limit (LSL)** and the **Upper Specification Limit (USL),** of your product.

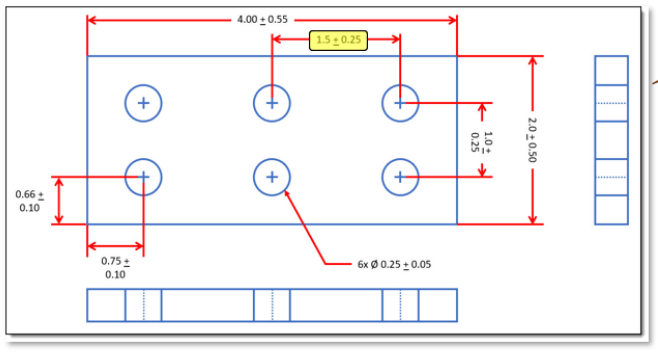
Capability Compares Process Performance v. Specifications

Now let's bring it all together to understand how your **process specifications** & **process performance** are used to calculate **Process Capability**.

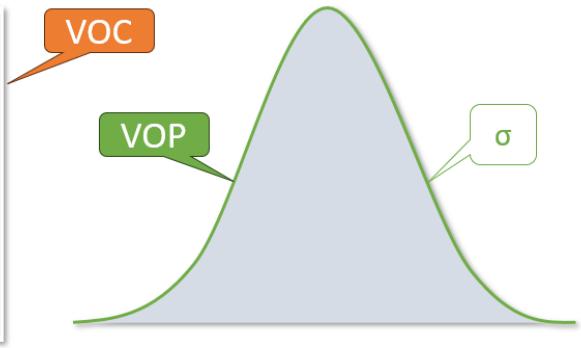
The concept of Process Capability can be summed up in the equation below:

$$Process\ Capability = \frac{Process\ Specification}{Process\ Performance} = \frac{Product\ Design\ Tolerance}{Natural\ Process\ Variation}$$

$$Process\ Capability = \frac{\boldsymbol{USL - LSL}}{\boldsymbol{6\sigma}}$$



Specifications

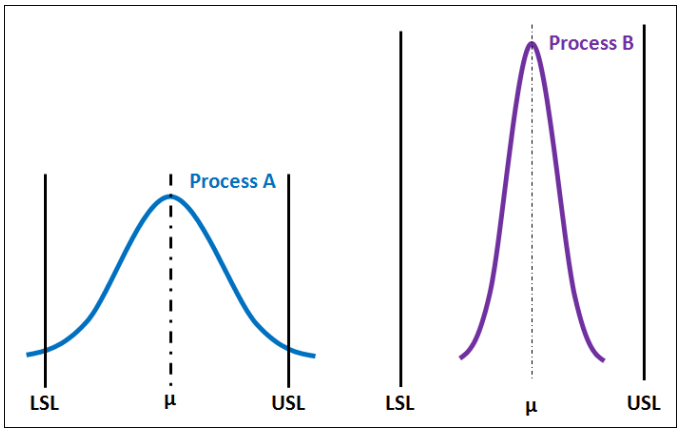


Process Variation

Example of Process Capability

So, if we take our histograms from the process above and add some fake specification limits to them, we can find out of process capability.

Note - These specifications *<u>are not the same</u>* as the control limits shown on the control chart above.



Remember, the histogram represents the natural variation of the process and the process specifications are shown as the Lower Specification Limit (LSL) & Upper Specification Limit (USL).

So let's say, in this example that the Lower Specification Limit is 2, and the Upper Specification Limit is 14, and the standard deviation of **Process A** is 3, and the standard deviation of **Process B** is 1.

$$Process\ Capability\ of\ \boldsymbol{A} = \frac{Process\ Specification}{Process\ Performance} = \frac{\boldsymbol{USL - LSL}}{\boldsymbol{6\sigma}} = \frac{\boldsymbol{14 - 2}}{6 * \boldsymbol{3}} = \frac{12}{\boldsymbol{18}} = 0.66$$

$$Process\ Capability\ of\ \boldsymbol{B} = \frac{Process\ Specification}{Process\ Performance} = \frac{\boldsymbol{USL - LSL}}{\boldsymbol{6\sigma}} = \frac{\boldsymbol{14 - 2}}{6 * \boldsymbol{1}} = \frac{12}{\boldsymbol{6}} = 2.00$$

Notice that in both of these calculations the **Process Specification** is identical (12 - 2 = 10).

What does change is the **Process Performance**, where Process A has 3 times the variation as Process B, and thus has less Capability.

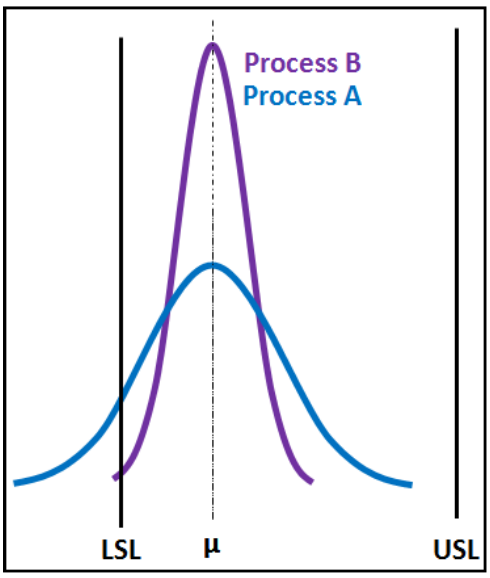
This reflects the image above, where the tails of Process A pass beyond this specification limits.

Another Example C_p Calculation - Shifted Limits

Let's now take a look at what happens to the process capability when the process specifications are shifted.

So the new **LSL is 4** and the new **USL is 16**.

You can see what that looks like on the histogram below, and the percentage of the population that is now out of specification (the tail of the distribution that's now past the LSL).



Now that the limits have shifted we can calculate Process Capability.

$$Process\ Capability\ of\ \boldsymbol{Process\ A} = C_p = \frac{Process\ Specification}{Process\ Performance} = \frac{\boldsymbol{USL - LSL}}{\boldsymbol{6\sigma}}$$

$$Process\ Capability\ of\ \boldsymbol{Process\ A} = C_p = \frac{16 - 4}{6 * 3} = \frac{12}{18} = 0.66$$

You'll notice here that C_p does not change with this shift.

This is because C_p does not take into account the center of your process.

I wanted to point this out before jumping into the details to let you know that the "Center" of your process becomes important when it comes to process capability.

It's one of the reasons that we have so many different Indices that can be used to calculate process capability.

I'll cover all of this below along with how to interpret your results.

Understanding the Prerequisites of a Process Capability Study - 5 Steps

Ok - this section is dedicated to the prerequisites of your process capability analysis and breaks them down into a 5 step checklist.

Step 1 - Identify the critical features that you want to study

We talked about this above, but I think it's worth repeating.

It isn't unusual for a product to have hundreds of features that are dimensioned & toleranced.

This usually makes it impossible to calculate process capability on all of these features.

So you must start by identifying the critical to quality product features and assess the process capability for only those features.

Step 2 - Ensure Your measurement System is Capable

Once you've identified a feature that you want to assess the capability of, the next step is to perform a gauge R&R to confirm that your measurement system is capable.

By capable I mean that the variation originating in your measurement system should be low relative to the specification of the feature that you're assessing. If a measurement system is a large source of variation, a capability study will be less meaningful and should be addressed prior to a capability study.

Step 3 - Ensure your Process is in Control

Again, this is a repeat but it's such a key concept that it's worth repeating.

For your process capability analysis to be valid, you must be analyzing a process that is "in control" and stable.

Remember, the overall goal of process capability analysis is to be able to make predictions about future products produces using that process and if the process is not in control, those predictions are meaningless.

If you think back to the chapter on SPC (Statistical Process Control), the most frequently used tool to ensure process control is a **Control Chart**. The good news is that the same data that is used to create a control chart can also be used to calculate Process Capability - and in fact this is often the case.

Remember, a stable process is one that is only experiencing normal cause variation, and no special cause variation.

A process that is not stable is experiencing special cause variation. If this is true for you, stop and eliminate the special cause variation before performing a capability study.

Step 4 - Ensure your data is Normally Distributed

This is the 2nd major requirement that must be verified before performing process capability analysis. **Your data must be normally distributed for the analysis to work out correctly.**

To confirm normality, you can use a Normal Probability Plot.

Now, if you're already a Pro at Process Capability Analysis you might be saying to yourself. . . non-normal data can still be used. This is true. But for now, let's assume the data must be normally distributed.

Step 5 - Understand How Your Variability is Calculated

Ok, so this is the first time we're going to cover this topic, but it's a super important one.

Below we're going to cover both C_p & P_p separately, but I'll show both equations here to make my point:

$$P_p = \frac{USL - LSL}{6\sigma} \qquad\qquad C_p = \frac{USL - LSL}{6\sigma}$$

Do these equations look similar??? YES!

The one distinction between these two equations is the way the Process Performance (s - Sample Standard Deviation) is calculated.

C_p only considered the process variation within a sub-group while P_p is calculated using the overall variation. Let's look at this calculation in a bit more detail for C_p.

When you're using data from your control chart (Step 2), you're calculating the sample standard deviation using the following equation.

$$\boldsymbol{s_{c_p}} = \frac{\boldsymbol{R}}{\boldsymbol{d_2}} \; (if\ using\ an\ \bar{X} - \ R\ Chart) \qquad or \qquad \boldsymbol{s_{c_p}} = \frac{\boldsymbol{s}}{\boldsymbol{c_4}} \; (if\ using\ an\ \bar{X} - \ S\ Chart)$$

This variation is sometimes considered the "short term" variation because it only considers "within Sub-group" variation & ignores the "between sub-group" variation that can be experienced by a process.

P_p & P_{pk} use a different standard deviation calculation that also takes into consideration the between-lot variation.

These indices calculate the sample standard deviation using the common root mean square calculation shown below.

$$\mathbf{s_{p_p}} = \sqrt{\frac{\sum(\mathbf{x} - \bar{\mathbf{x}})^2}{\mathbf{n} - \mathbf{1}}}$$

So, I say all of this to let you know that you must be cognizant of how you're calculating your standard deviation and what impact that has on your process capability analysis.

The Big 4 Process Capability Indices

As we discussed above, the concept of Process Capability can be summed up in the equation below:

$$Process\ Capability = \frac{Process\ Specification}{Process\ Performance} = \frac{Product\ Design\ Tolerance}{Natural\ Process\ Variation}$$

There are a handful of different indices (**C_p, C_{pk}, C_{pm} & C_r**) that can be used to calculate process capability.

The C_p Index

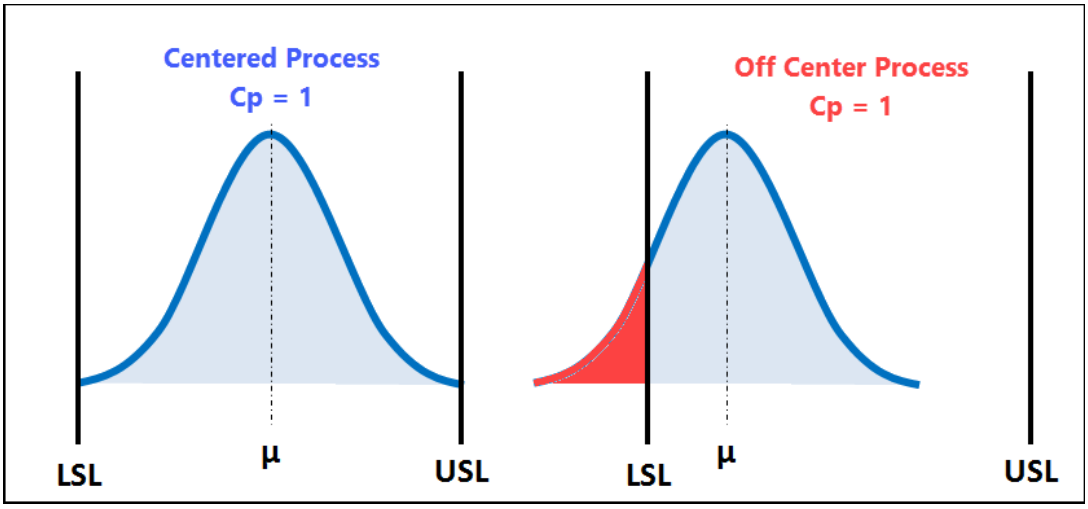
If you recall the example above, where the process was off center, our C_p calculation was still 1.0 however we had significantly more non-conforming product than before.

$$C_p = \frac{USL - LSL}{6\sigma}\ or\ \frac{USL - LSL}{6s}$$

In the example above, I'm using sigma (σ) to describe the standard deviation of the population, and s to represent the standard deviation of a sample. Use *sigma* when you've got it, and use s when you're working with a sample.

This is a good example of why C_p is often described **as a measure of the potential of the process** as it represents what the process might be able to achieve if the process was centered.

Remember this is because C_p only takes variability into account and ignores the location of the central value of the process.



While C_p only considered the variation within your process, C_{pk} takes into account both the variability & and centrality of the process and therefore reflects the actual performance of the process.

The C_{pk} Index

Ok, so let's take a look at the C_{pk} calculation for the example shown above to demonstrate how this index accounts a process that is not centered between the specification limits.

$$C_{pk} = Min(C_{p,Lower}, C_{p,Upper}) = Min\left(\frac{USL - \bar{x}}{3s}, \frac{\bar{x} - LSL}{3s}\right)$$

To break down this equation into terms, C_{pk} calculates the $C_{p,Upper}$ and $C_{p,Lower}$ and the returns the minimum value between these two.

You can see this in the image below, where the $C_{p,Upper}$ is the lower of the two values because the sample average is closer to the USL, and therefore more product falls outside of the USL.

This index penalizes any process that is not centered, and is a more accurate reflection of the process than Cp because it takes centeredness into account.

So let's **repeat our calculations** from above, where the **USL = 14, LSL = 2**, the sample standard deviation **(s) is 2,** and the process is centered with a **sample average ($\bar{x}$) of 8**.

$$C_{pk} = Min(C_{p,Lower}, C_{p,Upper}) = Min\left(\frac{USL - \bar{x}}{3s}, \frac{\bar{x} - LSL}{3s}\right) = Min\left(\frac{14 - 8}{3 * 2}, \frac{8 - 2}{3 * 2}\right)$$

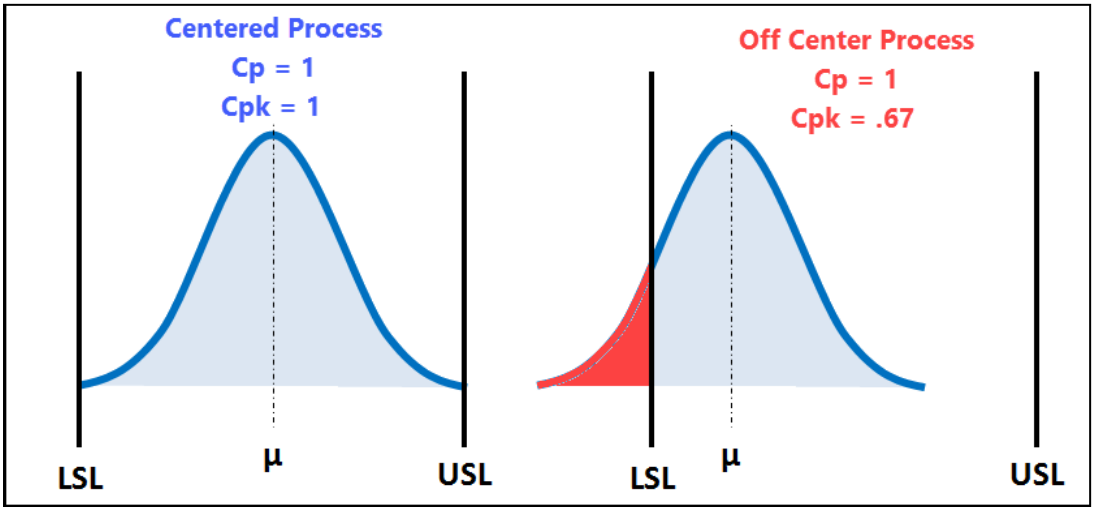
$$Centered\ Process: \qquad \boldsymbol{C_{pk}} = Min\left(\frac{6}{6}, \frac{6}{6}\right) == Min(1, 1) = \mathbf{1}$$

In this situation the C_{pk} is equal to C_p at 1.

Now let's see what happens to C_{pk} when the processes average is off-center after being shifted to 10.

$$\boldsymbol{Off\ Center}: \ \boldsymbol{C_{pk}} = Min\left(\frac{14 - 10}{3 * 2}, \frac{10 - 2}{3 * 2}\right) = Min\left(\frac{4}{6}, \frac{8}{6}\right) = Min(\,0.66\,, 1.33\,) = \mathbf{0.67}$$

Now that the process is off center, the Cpk falls to 0.67, and as you can see in the image below, this shifted process is now producing significantly more product that is outside of the specification limit.



The C_{pm} Index

C_{pm} is referred to as the Taguchi index, and (similar to C_{pk}) is designed to penalize processes that are not centered around the process Target Value (T).

Genichi Taguchi taught believed that negative cost impacts began the occur the moment your product deviates from the nominal or target dimension (even if it's within the tolerance).

The process Target Value (T) is generally the mid-point between the specification limits.

$$C_{pm} = \frac{USL - LSL}{6\sqrt{s^2 + (\bar{x} - T)^2}}$$

Ok, let's do an example of Cpk using the off centered process above. Just as a reminder our proceses sample average is 10 and a Process Target Value of 8.

$$\boldsymbol{C_{pm}} = \frac{USL - LSL}{6\sqrt{s^2 + (\bar{x} - T)^2}} = \frac{14 - 2}{6\sqrt{2^2 + (10 - 8)^2}} = \frac{12}{6\sqrt{4 + 4}} = \frac{12}{6 * 2.83} = \mathbf{0.707}$$

The C_r Index

C_r (The capability ratio) is the **inverse of C_p** and it represents the **percentage** of the **design specification** that is consumed by the process performance (variation).

$$C_r = \frac{1}{C_p} = \frac{6\sigma}{USL - LSL}$$

With the capability ratio, smaller is better. For example, a value of 0.10 would be interpretted that 10% of the design specification is consumed by the process variation. This would be ideal and would be equal to a Cp value of 10.0, which would be tremendous.

The 2 Must know Process Performance Indices

If you glance below at the P_p or P_{pk} equations, you'll notice they look very similar to the C_p & C_{pk} calculations.

$$P_p = \frac{USL - LSL}{6\boldsymbol{\sigma_{p_p}}} \qquad C_p = \frac{USL - LSL}{6\boldsymbol{\sigma_{c_p}}}$$

$$P_{pk} = Min\left(\frac{USL - \bar{x}}{3\boldsymbol{\sigma_{p_p}}}, \frac{\bar{x} - LSL}{3\boldsymbol{\sigma_{p_p}}}\right) \qquad C_{pk} = Min\left(\frac{USL - \bar{x}}{3\boldsymbol{\sigma_{c_p}}}, \frac{\bar{x} - LSL}{3\boldsymbol{\sigma_{c_p}}}\right)$$

But they have one important difference, and that is ***how the standard deviation is calculated*** *(Shown in red above)***.**

As we discussed above, C_p & C_{pk}, generally get their standard deviation from a control chart that only considers within-lot variation (equation shown below).

$$\boldsymbol{s_{c_p}} = \frac{\boldsymbol{R}}{\boldsymbol{d_2}} \text{ when using an } \bar{X} - \text{ R Chart}$$

This is sometimes considered the short-term standard deviation calculation because it ignores the between-lot variation that can be experienced by a process. This is also called the pooled standard deviation.

P_p & P_{pk} use a different standard deviation calculation that also takes into consideration the between-lot variation. These indices calculate the standard deviation using the common root mean square calculation:

$$\mathbf{s_{p_p}} = \sqrt{\frac{\sum(\mathbf{x} - \bar{\mathbf{x}})^2}{\mathbf{n} - \mathbf{1}}}$$

Note: The above calculation is for the *sample standard deviation as we generally are dealing with sample data for process capability analysis*. If the population standard deviation (Sigma - σ) is known, you can use that as well.

When we perform this calculation of the sample standard deviation it often includes data points that are associated with special cause variation.

This is why experts often say that P_p & P_{pk} does not require process stability as a prerequisite.

This decision to include "special cause variation" data points within your calculation of the sample standard deviation generally leads to a conservative estimate of the sample standard deviation.

However, your confidence in your P_{pk} calculation will be low if your process is not currently stable; and thus your ability to make predictions is low.

Normality is definitely still required for P_p & P_{pk} calculations to be valid.

With that, let's jump into each index.

The P_p Index

P_p is like C_p in that it only assesses the variability of your process & assumes the process is centered. It does not take the centrality of your process into account.

Similarly, if the process is not centered, then P_p becomes less representative of your actual performance and becomes more representative of the "Potential" of the process if it were centered.

$$P_p = \frac{USL - LSL}{6\boldsymbol{\sigma_{p_p}}}$$

The P_{pk} Index

As we discussed above, P_{pk} is calculated much like C_{pk} in that it considers both the central tendency of your process & the variability.

It's primary difference, as described above, is the way the standard deviation is calculated, which is why P_{pk} is often described as a more conservative estimate of long-term performance.

$$P_{pk} = Min\left(\frac{USL - \bar{x}}{3\boldsymbol{\sigma_{p_p}}}, \frac{\bar{x} - LSL}{3\boldsymbol{\sigma_{p_p}}}\right)$$

Let's say, for example, that you had a product with a dimension and tolerance of 12.0" $\pm$ 1.0".

You take 60 samples and you measure an average value of 12.2", and a standard deviation of 0.25".

Calculate the P_{pk} of your process.

$$P_{pk} = Min\left(\frac{13.0 - 12.2}{3 * 0.25}, \frac{12.2 - 11}{3 * 0.25}\right)$$

$$P_{pk} = Min(1.06, 1.60) = 1.06$$

Short Term versus Long Term Capability and the Sigma Shift

Over time, six sigma practitioners have recognized that the long-term performance of a process is not as good as the short-term performance of a process.

Over the long term, a process will like experience new sources of variation, or existing sources of variation can grow in magnitude. Long term variation is always greater than short term variation.

Additionally, the center of a process can drift over time, due to similar sources that impact the variation.

Sigma Level	% Conforming	DPMO	C_{pk}
1σ	68.3%	317,300	0.33
2σ	95.5%	45,500	0.67
3σ	99.7%	2,700	1.00
4σ	99.9937%	63	1.33
5σ	99.999943%	0.53	1.67
6σ	99.9999998%	0.002	2.00

Short-Term Quality Estimates based on Sigma Level or Cpk value

The differences between short-term performance and long-term performance was determined empirically. Meaning that it's based on experience and not necessarily any sort of science or data.

The ACTUAL difference between short- and long-term performance is based heavily on your process along with how much data you used to estimate your process capability.

A simple, mature, stable process will experience much less variation over time, and the 1.5 sigma shift might be much less. A complex, new, unstable process might experience more variation, and the long-term sigma shift might be much more.

Sigma Level	% Conforming	DPMO	C_{pk}
1σ	30.23%	697,700	0.33
2σ	69.13%	308,700	0.67
3σ	93.32%	66,810	1.00
4σ	99.3790%	6,210	1.33
5σ	99.97670%	233	1.67
6σ	99.999660%	3.4	2.00

Long-Term Quality Estimates based on Sigma Level or Cpk value

How To Interpret and React to Your Capability Results

Ok, so there are generally 5 ways to react once you've assessed the capability of your process; let's go through them 1 by 1.

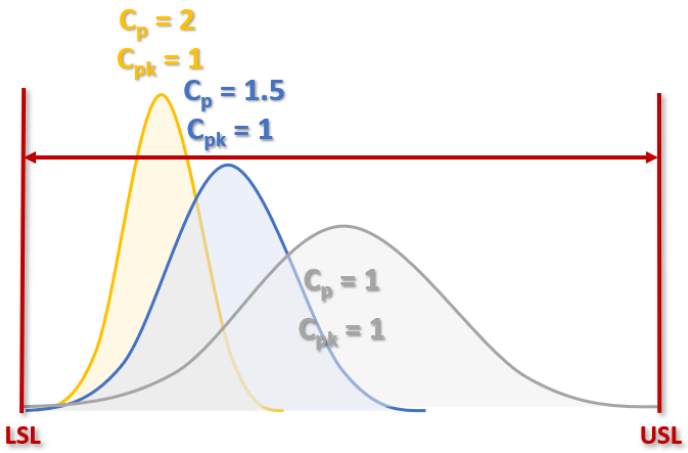
Before we go through these 5 reactions, let's refresh of the calculation for process capability:

$$\text{Cpk} = Min\left(\frac{USL - \overline{\boldsymbol{x}}}{3\boldsymbol{\sigma}}, \frac{\overline{\boldsymbol{x}} - LSL}{3\boldsymbol{\sigma}}\right)$$

The reason I want to review this calculation, is because if you're going to improve process capability, you're going to have to modify one of the variables used in that calculation.

These variables include the **mean** (x-bar), the **standard deviation (**σ), and the **specification limits** (USL or LSL).

Now what I want to do, is to show you three different processes, that all have the same C_{pk} calculation of 1.0.



What is different about these 3 processes, are their Cp calculations.

To understand the right reaction, you must compare your C_p against your C_{pk}.

If Your C_p >> C_{pk}, you should **center your process (reaction 2)**

If Your C_p ≈ C_{pk}, you should **reduce your variation (reaction 3)**

Before we review those reactions, let's talk about my favorite reaction 😊.

Reaction 1 - CELEBRATE! Cpk was great!

One potential result is that your process capability turns out to be fantastic! Congratulations, your process is awesome!

This is the DO-NOTHING reaction. Simply sit back and enjoy your capable process.

Reaction 2 - Adjust the Process to Center It Up

We've talked a lot about how important it is for your process to be centered, so that the tails of your normal distribution do not fall outside of your specification limits.

If Your $C_p >> C_{pk}$, you should **center your process (reaction 2)**

Remember Cp is the "potential" of your process, and if your Cp is bigger than your Cpk, it generally means that your process is not centered up.

The fastest way to better process capability is to **center up your process.**

Reaction 3 - Reduce the Variability associated with your process

If Your $C_p \approx C_{pk}$, you should **reduce your variation (reaction 3)**

In some cases, the right response to poor process capability is to reduce the normal variation associated with your process.

Oftentimes this is not a trivial task as it might require major redesigns to your process/equipment; however it might be your only option to better process capability.

Reaction 4 - Change the Process Specification

If you've considered all of the above, and there's nothing more you can do to improve your process, sometimes the right thing to do is to **change the Process Specification Limits.**

This usually means a conversation with your customer or your R&D team; and might include a change to the USL, LSL or both.

Reaction 5 - Assess the Risk Associated with Poor Process Capability

And now we're on to the absolute last resort. . . . .

If all else fails, and you can't reduce variability, or center your process, or change the process specifications, the only thing you can do is to accept the fact that your **Process Capability is poor**.

Now, **the decision to accept a process with poor capability should come with a risk assessment**.

This risk assessment should demonstrate you've considered all of the risks associated with your poor process capability and that those risks are acceptable.

Process Capability for Attribute Data

Up until this point, we've been measuring process capability using **variable data**.

However, oftentimes we monitor our process by counting up attribute data (pass/fail).

When dealing with attribute data, one thing to consider is how you are counting.

This means understanding the difference between a **defect and a defective**.

A "**defective"** is an **entire unit** that fails to meet specifications.

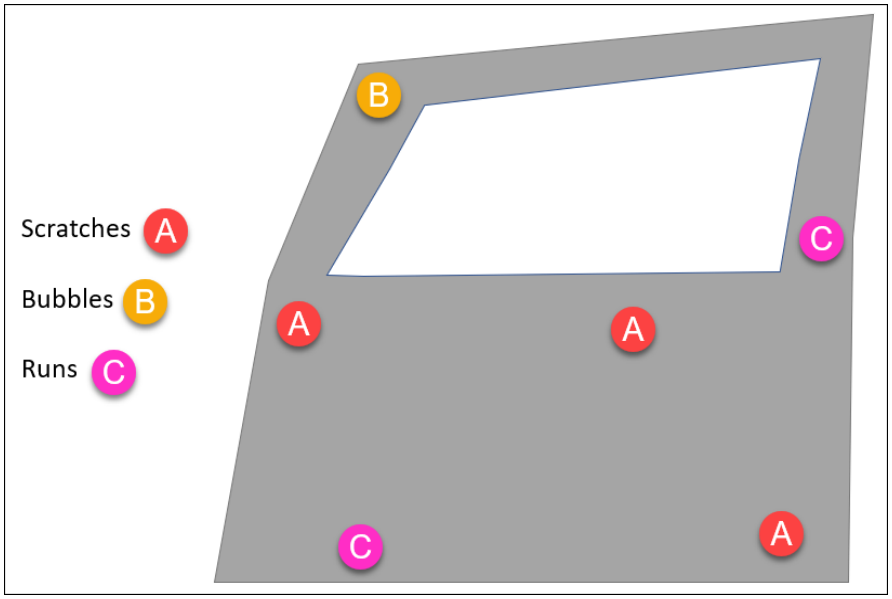
A "**defect"** is an **undesirable condition** within a unit.

A **defective unit can have multiple defects** associated with it, and any unit that has a **defect** on is considered **defective**.

The car door below illustrates both defects and defective items. This single unit has multiple defects on it (scratches, bubbles, runs), and is therefore considered a defective unit.

You could count this as 6 defects, or as a single defective unit.

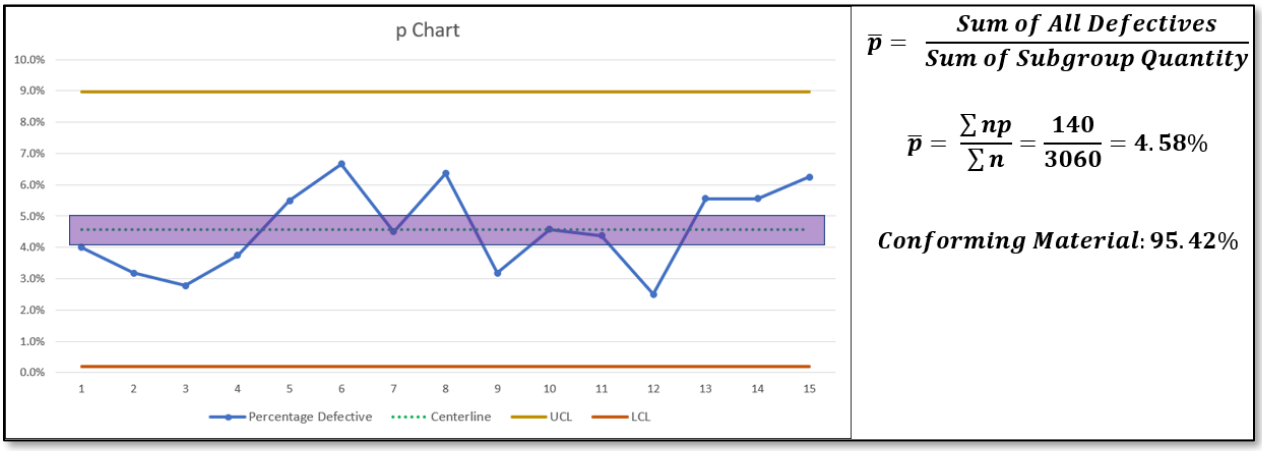
When measuring the quality of your process, you must decide how you're going to count.



For the purposes of this discussion, let's assume that you're counting defective items, and now let's talk about how you would measure process capability using this approach.

To teach this concept, I want to show you a control chart for attribute data called a P-chart, which we will cover in detail in chapter 27.

Below is a **P-chart, where we're tracking the percentage of defective units over time.**



The centerline of that control chart, highlighted in purple, is the **average defective rate.**

So, for this process, on average, 4.58% (percentage) of our defective items did not meet specifications.

That implies that 95.42% of your product is conforming to specifications.

Then, you can take that metric (95.4%) and go back to your process capability table (below) to equate your process performance back to a sigma level (2σ) and a C_{pk} level (0.67).

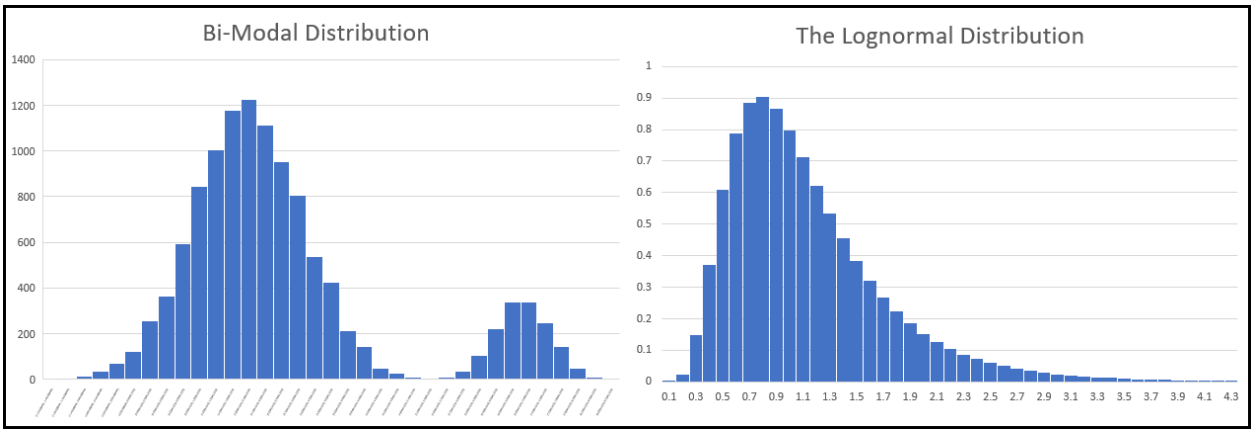
Sigma Level	% Conforming	DPMO	C_{pk}
1σ	68.3%	317,300	0.33
2σ	95.5%	45,500	0.67
3σ	99.7%	2,700	1.00
4σ	99.9937%	63	1.33
5σ	99.999943%	0.53	1.67
6σ	99.9999998%	0.002	2.00

Process Capability for Non-Normal Data

Above, we were discussing the requirements of a process capability study, and one of the expectations of your process, is that **the data from your process is normally distributed.**

Remember, to confirm normality, it's generally recommended to plot your data on a **normal probability plot**, as discussed in chapter 17.

For example, your data might be Bi-modally distributed, or it might have a heavy skew, such as a lognormal distribution.



This expectation for normality is not always met, but this doesn't mean that you can't calculate process capability.

If your data does not fit the normal distribution, you **must transform your data**!

Most software packages such as Minitab and others have various transformations that you can use to essentially transform your data into a normal distribution, these include:

- Box-Cox transformation
- Johnson Transformation

Don't Forget, if you're transforming your data, you must also **transform your specification limits** as well prior to performing any process capability calculations!

Process Performance Metrics

While process capability is a fantastic metric for the quality of your process, there are other metrics that you can use, these metrics include:

- Percent Defective
- Defective Parts Per Million (DPPM)
- Defects Per Unit (DPU)
- Defects Per Million Opportunities
- First Pass Yield
- Rolled Throughput Yield

These metrics are associated with discrete data that is simply reported in different forms (Percent Defective, DPU, DPMO).

For **discrete data** it means understanding the difference between a **defect and a defective**.

A "**defective"** is an **entire unit** that fails to meet specifications.

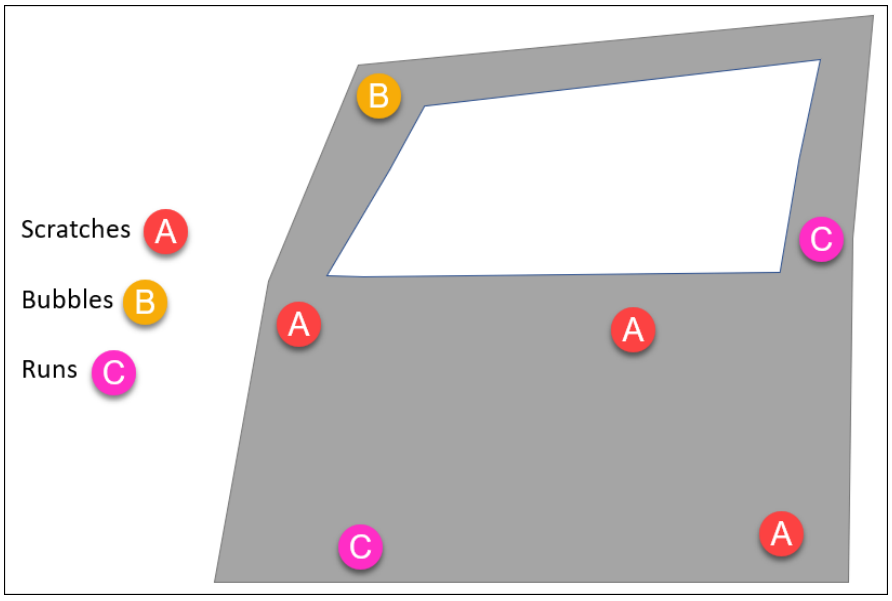
A "**defect"** is an **undesirable condition** within a unit.

A **defective unit can have multiple defects** associated with it, and any unit that has a **defect** on is considered **defective**.

The car door below illustrates both defects and defective items. This single unit has multiple defects on it (scratches, bubbles, runs), and is therefore considered a defective unit.

You could count this as 6 defects, or as a single defective unit.

When measuring the quality of your process, you must decide how you're going to count.



Percent Defective

One way to measure the health or the quality of your process is by counting the number of defective units.

So you can perform an inspection of products, looking for multiple defects, however whenever you find a defect, that single unit is considered a defective unit.

If you inspected 50 light bulbs, and found 7 defective items (regardless of their specific defect), you would calculate the percent of defective items as such:

$$\%\ \textbf{Defective} = \frac{\textbf{Sum of All Defectives}}{\textbf{Sum of Sample Size}}$$

$$\%\ \textbf{Defective} = \frac{7}{50} = 0.14\ \textbf{or}\ 14\%\ \textbf{Defective}$$

The percent defective items is a process performance metric that measures the proportion of faulty or substandard products within a production or service process.

This metric helps assess the quality and efficiency of a process, with lower percentages indicating better performance and quality control.

Defective Parts Per Million (DPPM)

The Defective Parts Per Million (DPPM) is a process performance metric used to quantify the level of defects in a production or manufacturing process.

It provides a more precise measure than the traditional percentage of defects, as it calculates the number of defective units per one million (1,000,000) units produced.

To calculate DPPM, follow these steps:

1. Count the number of defective items or parts produced within a given period.
2. Calculate the total number of units or parts produced in the same period.
3. Divide the number of defects by the total production, then multiply by one million

The formula for DPPM is as follows:

$$\textbf{Defective Parts Per Million} = \%\ \textbf{Defective} * 1,000,000$$

A lower DPPM value indicates better quality control, as it represents fewer defective parts per million units produced. In contrast, a higher DPPM indicates a lower level of process quality and a greater need for improvement.

DPPM is a valuable tool for identifying areas in a manufacturing process that require attention and improvement. **It allows for precise comparisons across different production runs, lines, or facilities.**

Defects per Unit (DPU)

This one is fairly straightforward. How many units did you make? How many defects were there?

It's important to note the difference between a "defect" and a "defective." A defect is a failure to meet specifications, and a single unit may have multiple defects associated with it.

To calculate the defects per unit, you must count the number of defects observed in an inspection, and divide by the total number of units that were included in that inspection.

$$\boldsymbol{DPU = \frac{Count\ of\ Defects\ Observed}{Total\ Number\ of\ Units\ Produced}}$$

Let's say, for example, you produce 100 motors in a given day, and throughout the day you found 136 defect conditions.

$$DPU = \frac{136}{100} = 1.36\ defects\ per\ unit$$

Notice here that there can be multiple defects per unit inspected, which can be common when you have a very complex product.

Defects Per Million Opportunities

Defects per Million Opportunities differs from DPU in that this metrics takes in consideration the ratio between how often a defect occurs and the potential for that defect to occur.

Defects per million opportunities requires you to quantify the number of possible defects that might be observed in each inspection.

Defects per million opportunities will allow us to estimate the error rate of our process, and allows you to make a fair, apples-to-apples comparison of one process against another.

$$DPMO = \frac{\boldsymbol{Count\ of\ the\ Number\ of\ Defects}}{\boldsymbol{Units\ Produced} * Number\ of\ Defect\ Opportunities\ per\ Part} * 1{,}000{,}000$$

You can simplify this by combining the count of defects and units produced into simply DPU.

$$DPMO = \frac{\boldsymbol{DPU}}{Number\ of\ Defect\ Opportunities\ per\ Part} * 1{,}000{,}000$$

For example, we can build off our example of motor inspections, where we inspected 100 motors, and found 136 defects.

Let's say, for example, that each motor has 500 possible defects, we can now calculate DPMO.

$$DPMO = \frac{\mathbf{136}}{\mathbf{100} * 500} * 1{,}000{,}000 = 2{,}720\ Defects\ Per\ Million\ Opportunities$$

The interpretation of this result is that if we produced 1,000,000 motors, we would expect to find 2,720 defects.

First Pass Yield

First pass yield is a quality metric that is used when **analyzing a series of connected processes**.

FPY is calculated as the good parts that exit the entire process divided by the total number of parts that started the process.

To find the number of good parts that exit the process, we must remove the scrap and rework.

$$\boldsymbol{First\ Pass\ Yield\ (FPY) = \frac{Total\ parts\ Produced\ - Scrap\ - Rework}{Total\ Parts\ Produced}}$$

So if you produced a work order of 1,000 pieces (lot size), and you count up 47 units scrapped, and 36 units reworked, the FPY would be calculated as:

$$\boldsymbol{First\ Pass\ Yield\ (FPY) = \frac{1,000 - 36 - 47}{1,000} = \frac{917}{1,000} = 91.7\%}$$

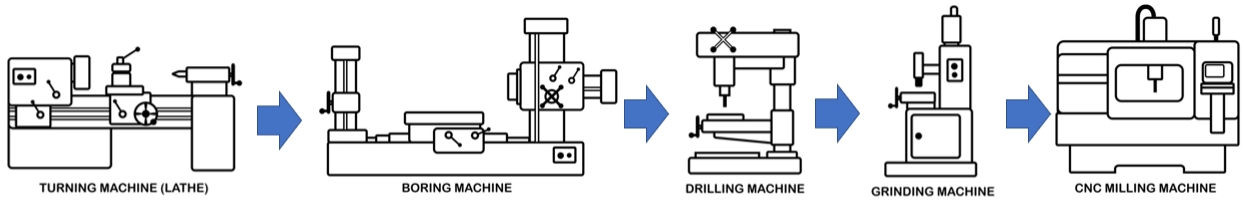
Rolled Throughput Yield

Rolled Throughput Yield is the probability that a **multi-step process** will produce a **defect free unit**.

RTY is calculated as the **product of yields** for each step in the process.

$$RTY = Y_1 * Y_2 * Y_3 * Y_4 \ldots * Y_n$$

Let's say for example, we had a 5-step process below, where each step had a 98% yield.



$$\boldsymbol{RTY = 98\% * 98\% * 98\% * 98\% * 98\% = 90.4\%}$$

There are two ways to interpret this value.

If 100 parts are fed into the process -> **90 good parts will be produced.**

If the customer orders 100 parts -> **111 parts should be fed into the process.**