

## Glossary

**2-level design:** an experiment where all factors are set at one of two levels, denoted as low and high (-1, 1 or 1, 2)

**2-tailed test:** also known as a two-sided test, it is a hypothesis test with a two-sided alternative hypothesis. That is, one could possibly err on either side of the center.

**3-level design:** an experiment where all factors are set at one of three levels, denoted as low, medium, and high (-1, 0, 1 or 1, 2, 3).

**5 Whys:** A tool that can be used to identify the root cause of an issue by asking the question “Why?” five times.

**5S Workplace Organization (5S):** A 5-step process to achieve and maintain a lean, clean and well organized workplace. The 5 S’s are: Sort, Store, Shine, Standardize, and Sustain.

**7 Forms of Waste:** Seven categories of waste were identified by Taiichi Ohno. These include: production of defects, over processing due to rework or poor process design, overproduction by producing before there is demand, the unnecessary movement of employees, unnecessary transportation of materials, holding more inventory than the absolute minimum, and waiting or delays which interrupt flow.

**Accelerated testing:** the testing of equipment in an environment so that the time will be shortened for failures to occur. For example, high temperature is often used to create failures sooner during a reliability test of electronic components.

**Acceptance region:** the region of values for which the null hypothesis is accepted.

**Accept quality level:** the maximum proportion of defective units in a sampling plan that can be considered satisfactory as the process average.

**Accuracy:** instrument accuracy is the difference between the observed average value of measurements and the master value.

**Active experimentation:** experiments are conducted where variable levels are changed to assess their impact on responses.

**Additive:** the total effect of several factors is equal to the sum of the individual factor effects.



**Affinity diagram:** a tool use to generate and organize ideas or issues into common themes. Once appropriately grouped, a single statement can summarize the idea or issues for each group. These statements can then be ranked in order of priority for further analysis or resolution.

**Aliasing:** when two factors or interaction terms are set at identical levels throughout the entire experiment (i.e., the two columns are 100% correlated).

**Alpha risk:** the probability of concluding the alternative hypothesis ( $H_a$ ) when the null hypothesis ( $H_0$ ) is true.

**Alternative hypothesis:** the hypothesis to be accepted if the null hypothesis is rejected. It is denoted by  $H_a$ .

**Analysis:** concerned with how it works, a study of the parts to determine the properties of the whole (system). Commonly associated with tolerance design.

**Analysis of goodness:** the ranking of fractional factorial experiments trails according to the level of a response. An attempt is then made to identify factors or combination of factors that potentially affect the response.

**Analysis of means (ANOM):** a statistical procedure to compare means of groups of common size to the grand mean.

**Analysis of variance (ANOVA):** a procedure for partitioning the total variation. It is often used to compare more than two population means.

**Analytic study:** a study in which action will be taken on a cause-and-effect system to improve performance of a product or a process in the future.

**Andon light:** A set of signal lights associated with a piece of equipment or a work cell that indicates it operational status.

**Appraiser variation:** the variation in average measurements of the same part between different appraisers using the same measuring instrument and method in a stable environment. AV is one of the common sources of measurement system variation that results from difference in operator skill or technique using the same measurement system.

**Arithmetic mean:** the balance point in a set of data that is calculated by summing the observed values in a set of data and then dividing by the number of values involved.

**ARIZ:** a Russian acronym for Algorithm for Solving Inventive Problems. It is a procedure for identifying conflicting requirements.



**Arrow diagram:** a diagram that shows cause and effect relationships using arrows.

**Assignable cause (of variation):** significant, identifiable change in a response which is caused by some specific variable from the cause and effect diagram.

**Attribute:** a characteristic that may take on only one value, e.g., 0 or 1.

**Attribute data:** Data that can be divided into various groups or categories on the basis of some non-numerical characteristics. This term is often used to describe data reflecting conformance or nonconformance to specifications, including observation of the presence or absence of some quality characteristic.

**Autonomous maintenance:** A subset of Total Productive Maintenance (TPM) which makes the worker the first line of maintenance.

**Average:** A measure of central tendency of a sample or the population. Synonymous with the term “mean.”

**Background variables:** variables which are of no experimental interest and are not held constant. Their effects are often assumed insignificant or negligible, or they are randomized to ensure that contamination of the primary response does not occur.

**Balanced design:** a 2-level experimental design is balanced if each factor is run the same number of times at the high and low level.

**Balanced scorecard:** a system of metrics that measures and displays all the different aspects of organizational performance.

**Bar chart:** a graphical method which depicts how data fall into different categories.

**Baseline performance:** snapshot of process performance before project improvements.

**Bathtub curve:** a curve used to describe the life cycle of a system/device as a function of usage. When the curve has its initial downward slope, the failure rate is decreasing with usage. This is the early-life region where manufacturing problems are typically encountered. The failure rate is constant during the flat part of the curve. When the curve begins sloping upward, the failure rate is increasing with usage. This region describes the wear-out of a product.

**Bell curve:** A bell shaped curve that graphically describes the probability distribution for “normal” data. Also known as a normal distribution.



**Benchmarking:** a structured approach to identifying a world-class process, then gathering relevant information and applying it within your own organization to improve a similar process.

**Benefit/Effort matrix:** sometimes called an Impact/Effort matrix, enables you to determine which project ideas have the most value.

**Best-subsets approach:** evaluates either all possible regression models for a given set of independent variables or the best subsets of models for a given number of independent variables.

**Beta risk:** the probability of concluding the null hypothesis ( $H_0$ ) when the alternative ( $H_a$ ) is true.

**Bias (in measurement):** systematic error which leads to a difference between the average result of a population of measurements and the true, accepted value of the quantity being measured.

**Bimodal:** Having two modes

**Bimodal distribution:** a distribution with two concentrations of data (two modes).

**Binomial distribution:** The probability distribution which describes a process or experiment in which each event has only two possible outcomes: pass/fail, defective/not defective, presence/absence, etc.

**Binomial experiment:** an experiment with a fixed number of independent trials. Each outcome falls into exactly one of two categories.

**Bivariate data:** data arranged in pairs.

**Bivariate normal distribution:** with paired data, for any fixed value of one variable, the values of the other variable have a normal distribution.

**Blocking variable:** a variable (factor) which cannot be randomized. The experiment is usually run in blocks for each level of the blocking variable and randomization is performed within blocks.

**Blocks:** groups of experimental units treated in a similar way in an experimental design; usually defined by background variables.

**Boldness:** the term used to describe the choosing of the magnitude of the variable levels to use within a response surface design. The concept suggests that the magnitudes of



variables should be large enough to capture the minimum and maximum responses of the process under test.

**Bootstrapping:** a resampling technique that provides a simple but effective method to describe the uncertainty associated with a summary statement without concern about details of complexity of the chosen summary or exact distribution from which the data are calculated.

**Bottleneck:** The rate limiting step of a process when demand exceeds capacity. Continually improving the performance of the bottleneck should improve the output of the whole operation until it has been elevated to the point where it is no longer the rate limiting step.

**Bottom line:** the final profit or loss that a company experiences at the end of a given period of time.

**Box-Behnken design:** a 3-level design used for quantitative factors and designed to estimate all linear, quadratic, and 2-way interaction effects.

**Boxplot:** a graphic method of showing the first, and third, quartiles as well as the median (2<sup>nd</sup> quartile) of the sample data.

**Brainstorming:** a group activity which generates a list of possible factors and levels, and the method by which the results may be evaluated.

**Buffer resources:** a means of meeting customer demand when customer ordering patterns, or takt times, vary.

**Calibration:** adjusting an instrument using a reference standard to reduce the difference between the average reading of the instrument and the “true” value of the standard being measured, i.e., to reduce measurement bias.

**Capability of a process:** A measure of the quality of a process that is derived by comparing the voice of the process (a distribution) to the voice of the customer (specification limits). Capability measures include sigma level, sigma capability, Cp, Cpk, and defects per million (dpm).

**Catchball:** a give-and-take activity performed between different levels of the organization to make sure that critical information on goals and objectives as well as feedback is passed back and forth.

**Categorical variable:** a variable whose results consist of a tally of the number of items in each category.



**Causality:** the assertion that changes to an input factor will directly result in a specified change in an output.

**Causal loop:** a tool for documenting the cause and effect relationship between measurable indices, states, etc.

**Causal relationship:** a relationship where variation in one variable causes variation in another. Statistical tests can show a relationship between variables but not that it is causal. Other factors might be involved in the relationship.

**Cause and effect diagram (CE):** A tool that is used to categorize, display, and examine potential causes or contributing factors related to a specific observed condition. This tool is also known as a fishbone diagram.

**C chart:** charts which display the number of defects per sample.

**Cellular operations:** method of designing the physical layout of a process in a way that minimizes space, time, equipment, and inventory. Also called cellular manufacturing or cellular layout.

**Center line:** the line on a statistical process control chart which represents the characteristic's central tendency.

**Center points:** experimental runs with all factor levels set halfway between the low and high settings.

**Central composite design:** a 3-level design that starts with a 2-level fractional factorial and some center points. If needed, axial points can be tested to complete quadratic terms. Used typically for quantitative factors and designed to estimate all linear effects plus desired quadratics and 2-way interactions.

**Central limit theorem:** theorem stating that sample means tend to be normally distributed.

**Central tendency:** A measure of the point about which a group of values is clustered. Examples include mean, median, and mode.

**Champion:** manager or business leader who spearheads the selection, implementation, and completion of Lean Six Sigma projects.

**Change:** an event that causes a problem to commence.

**Changeover:** All of the activities associated with switching the materials, operating settings or tooling on a piece of equipment so that it can produce a different part of



perform a different task. Changeover time is usually defined as the time elapse between the production of the last good part of one batch until the first good part is produced of the next batch.

**Characteristic:** a process output which can be measured and monitored for control and capability.

**Characteristic life:** a parameter that is contained in the Weibull distribution. In a reliability test, the value of this parameter equates to the usage when 63.2% of the devices will fail.

**Characteristic matrix:** a tool that identifies the relationships between technical characteristics.

**Chi-square distribution:** the distribution of chi-square statistics.

**Chi-square:** the test statistic used when testing the null hypothesis of independence in a contingency table or when testing the null hypothesis of a set of data following a prescribed distribution.

**Check sheet:** a data gathering sheet which categorizes data into time slots, machines, departments, shifts, employee, etc.

**Check solution:** to test a solution for effectiveness and practicality.

**Chronic problem:** a description of the situation where a process SPC chart may be in control/predictable; however, the overall response is not satisfactory (i.e., common causes yield an unsatisfactory response). For example, a manufacturing process has a consistent “yield” over time; however, the average number is not satisfactory.

**Chunk variable:** a variable developed by forming blocks of a certain combination of background variables.

**Classical methods:** statistical experimental design thoughts and processes as originally developed by Fisher and others as early as the 1920’s. Uses ANOVA as the primary tool, along with orthogonal designs such as fractional factorials, latin squares, Plackett-Burman, Box-Behnken, central composite, and D-optimal.

**Class variables:** factors that have discrete levels.

**Clean data:** data that is accurate and free of correlation between input factors.

**Cluster sampling:** population is divided into sections and a sampling of sections is randomly selected.



**Coded levels:** regression analysis of factorial or response surface data can be performed where the levels are described in the natural levels of the factors (e.g., 5.5 and 4.5 V) or the coded levels of the factors (e.g., -1 and +1).

**Coefficient of correlation:** measures the strength of the linear relationship between two numerical values.

**Coefficient of determination ( $R^2$ ):** the amount of variation in y that is explained by the regression line.

**Coefficient of multiple determination:** represents the proportion of the variation in the dependent variable Y that is explained by the set of independent variables selected.

**Coefficient of determination ( $R^2$ ):** the square of the sample correlation coefficient; it represents the strength of a model.  $(1 - R^2) * 100\%$  is the percentage of noise in the data not accounted for by the model.

**Coefficient of variation:** the ratio of the standard deviation to the mean. It is a standardized method of looking at variation.

**Collectively exhaustive events:** a set of events such that one of the events in the set must occur.

**Collinearity of the independent variables:** refers to situations in which one or more of the independent variables are highly correlated with each other.

**Common cause variation:** The sources of variability in a process that are truly random and occur naturally as an inherent part of the process.

**Completely randomized design:** in analysis of variance, each element is given the same chance of belonging to the different categories or treatments.

**Composite design:** a design for evaluating nonlinear factor effects that is constructed by adding selected factor combinations to two-level factorial designs.

**Concept development:** developing and selecting design concepts for systems, subsystems, components, or processes. Also known as system design.

**Concurrent engineering:** an approach to the development of new products where the product and all its associated processes, such as manufacturing, distribution, and service, are all developed in parallel.





**Confidence:** likelihood of not making a Type I error ( $1 - \alpha$ ).

**Confidence interval:** range within which a parameter of a population (e.g., mean, standard deviation, etc.) may be expected to fall, on the basis of measurement, with some specified confidence interval or confidence coefficient.

**Confidence limits:** the upper and lower boundaries of a confidence interval.

**Confirmation run:** a test of the preferred combination which is determined from the analysis of the experiment. It is used to verify the conclusions of the analysis.

**Conflicting requirements:** customer or internal requirements that cannot be met simultaneously because they contradict one another.

**Confounded:** two factors are considered to be confounded when their test profiles contain the same pattern of test settings, thus making it impossible to evaluate the two factors independently.

**Confounded effects:** the average effect of a factor, or a differential effect between factors (interactions), combined indistinguishably with the effects of other factor(s), block factor(s), or interaction(s).

**Confounding factor:** an independent variable (in addition to the one under test) that has a systematic influence on the dependent variable.

**Consumers risk:** the probability of accepting a lot when, in fact, the lot should have been rejected (see Beta risk).

**Contingency diagram:** a tool that identifies actions to take if a solution should fail.

**Contingency table:** a table of observed frequencies where the rows correspond to one variable of classification and the columns correspond to another variable of classification.

**Continuous flow:** the ideal state characterized by the ability to replenish a single work unit that has been “pulled” downstream. In practice, continuous flow is synonymous with just-in-time (JIT), which ensures that both internal and external customers receive only what is needed, just when it is needed, and in the exact amounts needed.

**Continuous variable:** a variable that is derived from measurement data and is considered to have an infinite number of possible values. That is, it is measured on a continuous scale.

**Control (of process):** a process is said to be in a state of statistical control if the process exhibits only random variation (as opposed to systematic variations and/or variations



with known sources). When monitoring control with control charts, a state of control is exhibited when all points remain between set control limits without any abnormal (non-random) patterns.

**Control chart:** the basic tool of statistical process control. It consists of a run chart, together with statistically determined upper and lower control limits and a centerline.

**Control factors:** factors that can be assigned at specific levels, set by those designing the product or process.

**Control group:** a group of subjects matched with the experimental group on all relevant factors except the experimental manipulation. For example, a placebo group (who do not take a particular drug) could be used as a control group for a drug group (who do) to examine the effect of the drug on performance.

**Control limits:** upper and lower bounds in a control chart that are determined by the process itself. They can be used to detect special causes of variation. They are usually set at  $\pm 3$  standard deviations from the centerline.

**Control plan:** a listing of controls used for process control. The control plan will typically include: part/process number, equipment type, characteristics, product/process specification, tolerance, measurement technique, sample size and frequency, control method, and reaction plan.

**Co-Relations diagram:** a tool that shows how technical characteristics either contradict or complement each other in meeting customer requirements.

**Co-Relations matrix:** a tool that summarizes the results of co-relations diagrams.

**Core team:** team that configures, drives, reviews, and is responsible for the Lean Six Sigma initiative.

**Correlation:** the degree to which the scores (from a set of subjects) on two variables correlate. That is, the extent to which a variation in the scores on one variable results in a corresponding variation in the scores on the second variable.

**Correlation coefficient (r):** a measure of the linear relationship between two random variables.

**Controllable factors:** factors the experimenter has control of during all phases, i.e., experimental, production, and operational phases.



**Cost of quality:** an analysis of financial indicators related to quality, including costs of conformance, including prevention and appraisal costs, and costs of nonconformance, including internal and external losses as a result of unacceptable quality (nonconformity).

**Covariates:** things which change during an experiment which had not been planned to change, such as temperature or humidity. Randomize the test order to alleviate this problem. Record the value of the covariate for possible use in regression analysis.

**$C_p$ :** during process capability studies,  $C_p$  is a capability index which shows the process capability potential but does not consider how centered the process is.  $C_p$  may range in

value from 0 to infinity, with a large value indicating greater potential capability. A value of 1.33 or greater is usually desirable.

**$C_{pk}$ :** during process capability studies,  $C_{pk}$  is an index used to compare the natural tolerances of a process with the specification limits.  $C_{pk}$  has a value equal to  $C_p$  if the process is centered on the nominal; if  $C_{pk}$  is negative, the process mean is outside the specification limits; if  $C_{pk}$  is between 0 and 1 then the natural tolerances of the process fall outside the spec limits. If  $C_{pk}$  is larger than 1, the natural tolerances fall completely within the spec limits. A value of 1.33 or greater is usually desired.

**Creative design:** a design that is the product of one's ingenuity and uses design principles in a unique way.

**Critical region:** the area under a curve containing the values that lead to rejection of the null hypothesis.

**Critical success factors:** factors that are necessary to the success of Lean Six Sigma deployment.

**Critical value:** value separating the critical region from the values of the test statistic that would not lead to rejection of the null hypothesis.

**Critical X worksheet:** sorts significant x's from those that don't impact the process or Y of concern.

**Critical-To characteristic (CT):** any feature of a product or process that is important to the customer or the business.

**CT Flowdown:** the cause-and-effect chain of critical factors related to any process or product.



**Curvature:** refers to non-straight line behavior between one or more factors and the response. Curvature is usually expressed in mathematical terms involving the square or cube factor.

**Customer context table:** a tool for documenting the habits of a particular customer in detail.

**Customer demand:** the quantity of work units required by the customer.

**Customer segmentation table:** a tool for documenting the habits of different types of customers.

**Customer voice table:** a tool for documenting the verbal comments of customers concerning the requirements of a product, service, system, etc.

**Cycle time:** the elapsed time for a product to progress completely through the process. Cycle time begins from the start of the first step until the end of the last step.

**D-optimal design:** an experimental design in which the minimum number of runs is based on degrees of freedom needed to analyze the desired effects. Not necessarily orthogonal or balanced, it does minimize the correlation (confounding) between factors.

**Data:** the numbers or information collected in a project or experiment.

**Data collection sheet:** used to collect detailed information on potential x's and supply specific x – Y data for analysis.

**Decision tree:** a graphical decision-making tool that integrates for a defined problem both uncertainties and cost with the alternatives to decide on the “best” alternative.

**Defect:** departure of a quality characteristic from its acceptable level or state, i.e., the measured value of the characteristic is outside of specification. Also referred to as non-conformance to requirements.

**Defective unit:** a sample (part) which contains one or more defects, making the sample unacceptable for its intended, normal use.

**Defects per million (dpm):** a capability measure which describes the number of defective units per million units produced. When a “unit” is considered to be an “opportunity” for failure, this is sometimes referred to as dpmo (defects per million opportunities).

**Defining relationship:** a statement of one or more factor word equalities used to determine the aliasing structure in a fractional factorial design.



**Defining words:** factor word equalities in building a defining relationship.

**Degree of confidence:** probability that a population parameter is contained within a particular confidence interval.

**Degrees of freedom:** a parameter in the t, F, and chi square distributions. It is a measure of the amount of independent information available for estimating the population variance. It is the number of independent observations minus the number of parameters estimated.

**Delphi technique:** a method of predicting the future by surveying experts in the area of concern.

**Dependability:** the ability of a product/system/process to function during its target life, with performance which is within a window of acceptability as defined by customer expectations. Dependability is an overall customer-focused measure which encompasses all of the aspects of the product/system/subsystem/component quality, reliability, durability, and performance (QRDP).

**Dependent variable:** the variable measured by the experimenter and predicted to be influenced by (i.e., depend on) the independent variable.

**Deployment:** the act of spreading, coordinating, and driving many improvements throughout an organization.

**Deployment champion:** a leader who drives and oversees Lean Six Sigma deployment throughout an organization.

**Deployment infrastructure:** systems, supports, functions, and software required to proliferate Lean Six Sigma throughout an organization.

**Deployment planning:** process of defining and documenting how an organization will deploy Lean Six Sigma or some other initiative.

**Deployment tree:** a tree diagram that shows how an idea or concept is translated into greater and greater detail, to a set of actionable items.

**Descriptive statistics:** the methods used to summarize the key characteristics of known population data.

**Design a cushion:** a solution in which a design change reduces the effect.



**Design for Six Sigma (DFSS):** Design for Six Sigma represents a set of more advanced tools and methodologies for significantly reducing variation as early in the development cycle as possible. It applies to the design of new products and processes as well as their re-design.

**Design of experiments:** an organized method of collecting data which purposely changes the process inputs in order to observe and measure the corresponding changes in the process output. DOE is a method for developing powerful empirical models which approximate the true relationships between various process inputs and the output(s). Understanding these relationships allows us to improve the process performance characteristics.

**Design matrix:** an array representing the experimental settings. Usually contains values ranging from -1 to +1, but could be wider if using a CCD. The rows represent the runs and the columns represent the factors.

**Design problem:** a problem in which a design cannot meet all the requirements due to an inherent flaw.

**Deviation:** the difference between an observed value and the mean or average of all observed values.

**Direct cause:** in a cause and effect chain, the cause that is closest to the observed effect.

**Discrete data:** data resulting from either a finite number of possible values or a countable number of possible values.

**Discrete random variable:** a variable that is based on count data (number of defects, number of births, number of deaths, etc.). It supposedly has a countable number of possible outcomes.

**Discrimination:** the technological ability of the measurement system to adequately differentiate between values of a measured parameter.

**Dispersion:** the tendency of the values of the elements in a sample to differ from each other. Dispersion is commonly expressed in terms of the range of the sample or by the standard deviation.

**Distinction:** a description of what is “different” between two items that should be the same but aren’t.

**Distribution:** tendency of large numbers of observations to group themselves around some central value with a certain amount of variation or “scatter” on either side.



**Divide and conquer:** a strategy in which large problems are broken down into smaller problems, which individually are straightforward to solve.

**DMAIC:** a 5-phase strategy for process improvement: Define, Measure, Analyze, Improve and Control.

**Dot diagram:** a plot, using a number line centered at zero, of the estimated effects from a factorial study. Effects clustered near zero on the diagram cannot be distinguished from variation due to nuisance variables and should not be considered important.

**Dot frequency diagram:** a graph used to analyze data from a nested experimental pattern. The graph is designed to visually partition the variation in the data among the nested factors.

**Double-sided test:** a statistical consideration where, for example, an alternative hypothesis is that the mean of a population is not equal to a criterion value.

**DPMO:** When using the nonconformance rate calculation of defects per million opportunities (DPMO) one needs to first describe what the opportunities for defects are in the process (e.g., the number of components and solder joints when manufacturing printed circuit boards). Next the number of defects is periodically divided by the number of opportunities to determine the DPMO rate.

**Effect:** change observed in the response when a factor is varied from one level to another.

**Effect size:** the size of the difference between the means of two populations, in standard deviation units.

**Effects Table:** table of effects for all main effects and interactions (also called a response table, delta table or means table).

**Effects tree:** a tree diagram that shows the far-reaching effects of a proposed solution.

**Efficiency:** measure of the sensitivity of a nonparametric test in comparison to a corresponding parametric test.

**Eight wastes:** Lean concepts of Waiting, Overproduction, Rework, Motion, Processing, Inventory, Transportation, and Information.

**80/20 rule:** frequently related to Pareto charts, this rule states that 80% if any effect is caused by 20% of the input variables. Also known as the “vital few” versus the “trivial many.”





**Eliminate the conflict:** a design solution that uses creative design to meet all requirements equally, including those that contradicted each other.

**Entitlement:** Best performance possible based on process design, disregarding uncontrolled variation.

**Enumerative study:** a study in which action will be taken on the universe.

**EPIC solution matrix:** method of rating possible solutions according to the dimensions of Ease, Performance, Impact, and Cost.

**Error (experimental):** ambiguities during data analysis caused from such sources as measurement bias, random measurement error, process noise, etc.

**Error proofing (poka-yoke):** a series of techniques which are used to reduce the opportunity for errors to occur.

**Estimation:** an approach to making inferences about population parameters. This includes both point estimates and interval estimates (confidence intervals).

**Event:** a result of outcome of an experiment.

**Evolutionary operation (EVOP):** a strategy to run experiments on a process in operation. The EVOP studies were designed to be run by operators on a full-scale manufacturing process while they continued to produce output of satisfactory quality.

**Expectation pareto:** a pareto chart that uses expected gain or loss instead of number of occurrences as the measure of comparison.

**Expected frequency:** theoretical frequency for a cell of a contingency table or multinomial table.

**Expected value:** for a discrete random variable, the sum of the products obtained by multiplying each value of the random variable by the corresponding probability.

**Experiment:** an analytic study to provide a basis for action.

**Experimental design:** purposeful changes to the inputs (factors) to a process in order to observe corresponding changes in the outputs (responses).

**Experimental error:** variation in observations made under identical test conditions. Also called residual error. The amount of variation which cannot be attributed to the variables included in the experiment.



**Experimental pattern:** the arrangement of factor levels and experimental units in the design.

**Experimental unit:** the smallest division of material in an experiment such that any two units may receive different combinations of factors. Examples of experimental units are a part, a batch, one pound of material, an individual person, or a 10 square foot plot of ground.

**Explained deviation:** for one pair of values in a collection of bivariate data, the difference between the predicted y value and the mean of the y values.

**Explained variation:** the sum of squares of the explained deviations for all pairs of bivariate data in a sample.

**Exponential distribution:** a probability distribution mathematically described by an exponential function. Used to describe the probability that a product survives a length of time in service, under the assumption that the probability of a product failing in any small time interval is independent of time.

**External customer:** person or organization who purchases the final product or service.

**External changeover steps:** changeover activities that can be done while the process is running.

**External noise factors:** noise factors relating to the environment in which the product is used or distributed.

**F table:** provides a means for determining the significance of a factor to a specified level of confidence by comparing calculated F ratios to those from the F distribution. If the F ratio is greater than the table value, there is a significant effect.

**F-distribution:** distribution of F-statistics.

**F-ratio:** a ratio of two independent estimates of experimental error. If there isn't an effect, the ratio should be close to 1.

**Factor:** an input to a process which can be manipulated during experimentation. Also referred to as parameters and independent variables.

**Failure Mode and Effect Analysis (FMEA):** a tool which is used to systematically identify key modes of process or product failure, analyze risks, and prioritize improvement efforts.



**Failure rate:** the average number of failures per unit time. Used for assessing reliability of a product in service.

**FIFO:** First in, first out (FIFO) is a work-control method used to ensure that the oldest work (first in) is the first to be processed (first out).

**Find conflicts:** a strategy for finding pairs of conflicting requirements in designs.

**First pass yield:** the measure typically referred to as “yield.” The total number of parts that are accepted divided by the total number of parts that were started.

**First time quality:** a measurement indicator of process capability used in conjunction with other quality measurements, such as customer complaints, final audit, to develop specific improvement plans at the source and ultimately improve outgoing quality.

**First-time-yield (FTY):** extent to which a process produces its intended outcome right the first time.

**Fishbone diagram:** a tool for documenting possible causes of a problem.

**Five S's:** five disciplines that can transform a messy, cluttered, and inefficient process into one that is clean and streamlined. Stands for Sort, Store, Shine, Standardize, and Sustain.

**Fixed effects model:** experimental treatments are specifically selected by the researcher. Conclusions only apply to the factor levels considered in the analysis. Inferences are restricted to the experimental levels.

**Flow:** the movement of material or information. Businesses are successful to the extent that they are able to move material or information with as few disruptions as possible – preferably none.

**Flowchart:** a display of the various stages or steps in a process, using different types of symbols, to demonstrate the flow of product or service over time.

**Foldover design:** a way to obtain a resolution four design based on two designs of resolution three. Used when the combination of runs from a resolution III design differ substantially from their prediction and the experimenter desires to de-alias the 2-way interactions from their main effects.

**Force field diagram:** a tool that identifies forces that encourage implementation of a solution and forces that discourage implementation.



**Fractional factorials:** instead of using a full factorial, a subset or fraction of it can be used if the experimenter can assume some interactions will not occur.

**Frequency:** the number of times a score, a range of scores, or a category is obtained in a set of data.

**Frequency distribution:** the pattern or shape formed by the group of measurements in a distribution.

**Frequency table:** a list of categories of scores along with their corresponding frequencies.

**Full factorial:** all possible combinations of the factors and levels. Given  $k$  factors, all with two levels, there will be  $2^k$  runs. If the factors have three levels, then  $3^k$  runs would be needed.

**Gantt chart:** project management tool used to keep track of tasks, people, deadlines, and cost.

**Gemba:** The actual place where work is performed.

**Goodness-of-fit test:** test for how well some observed frequency distribution fits some theoretical distribution.

**Grand average:** overall average of data.

**Green Belt:** part-time Lean Six Sigma practitioner who completes fewer, less difficult projects than a Black Belt.

**Half effect:** the difference between the average of a response at a given level and the grand mean.

**Hard savings:** quantifiable financial savings that result directly from a Lean Six Sigma project.

**Heijunka or load leveling:** balancing the amount of work to be done (the load) during a day with the capacity to complete the work. A heijunka system distributes the work in proportions based on demand, factoring in volume and variety.

**Heijunka box:** a physical device used to level production volume and variety over a specified time period (usually one day). The box is divided into slots that represent pitch increments. The slots are loaded with kanbans that represent customer orders. The order in which kanbans are loaded into the box is determined based on volume and variety.



**Hidden factory:** rework built into a process as a “necessary” step. Prevents doing it right the first time.

**Histogram:** a bar chart that depicts the frequencies of numerical data.

**Homogeneity of variance:** the variances of the groups being contrasted are equal (as defined by statistical test of significant difference).

**Homoscedasticity:** the scores in a scatterplot are evenly distributed along and about a regression line. This is an assumption made in linear correlation.

**Hoshin planning:** practice of cascading top-level organizational strategies into increasingly lower levels, with performance metrics attached, and plans for improvement.

**Hypothesis:** a statement or claim that some population characteristic is true.

**Hypothesis test:** a procedure whereby one of two mutually exclusive and exhaustive statements about a population parameter is concluded. Information from a sample is used to infer something about a population from which the sample was drawn.

**Ideal manning levels:** determined by dividing the sum of your manual cycle times (time required to perform a job) by your Takt time (rate of customer demand).

**Ideal result (ARIZ):** a creative thinking tool in which one visualizes a design that meets all requirements and has absolutely no drawbacks (no cost, no weight, no failure modes, etc.)

**Immediate corrective action:** an action, usually interim, taken as quickly as practicable to contain a problem.

**Implementation plan:** documents all the actions required to implement project changes, as well as when the changes are to be done and by whom.

**Improvement cycle:** an adaptation of the scientific method (consisting of four phases – plan, do, study, act) used to increase a team’s knowledge about the product or process and to provide a systematic way of accomplishing change.

**Incomplete block design:** an experimental pattern in which the number of experimental units in a block is less than the number of combinations of factors and levels.

**Independent samples:** the values in one sample are not related to the values in another sample.



**Independent variable:** a variable chose by the experimenter for testing, predicted to influence the dependent variable.

**Inner array:** a Taguchi term used in parameter design to identify the combinations of controllable factors to be studied in a designed experiment.

**Input:** the sources of variation which may impact a process. These typically include categories such as materials, people, equipment, policies, procedures, methods, measurement, and environment.

**Input Process Output (IPO) Diagram:** a pictorial tool which is used to describe all of a process' inputs and its outputs.

**Instability:** Unnaturally large fluctuations in a pattern.

**Institutionalize:** to implement the principles of the solution throughout the company.

**Interaction:** 2 factors (input variables) are said to interact if one factor's effect on the response is dependent on the level of the other factor, i.e., the interdependence of two or more factors).

**Interim actions:** actions taken to contain a problem, which are meant to be superseded by a permanent solution.

**Internal changeover steps:** activities that must be done while a process is idle. Should be avoided or minimized whenever possible.

**Internal customer:** someone who needs the output of the previous process step to do one's job.

**Internal noise factors:** noise factors relating to product deterioration with age or use.

**Interval:** level of measurement of data: data can be arranged in order, and differences between data values are meaningful.

**Interval data:** data produced by the use of an interval scale. Parametric tests require interval data.

**Interval scale:** a scale of measurement where the interval between consecutive number is always the same. Most measuring devices, such as timers, thermometers, tape measures, employ interval scales.



**Inventive benchmarking:** convergent brainstorming technique where participants work their way to a specific solution through the use of established inventive principles and thinking by analogy.

**Inventive principles:** a set of 40 general design principles for creative design.

**Is – Is Not:** a tool in which a problem’s symptoms are thoroughly described, taking into account what has been observed and what has not been observed

**Just Do It Project:** a project for solving a problem when the solution is known and you just have to implement it.

**Just-in-time (JIT):** a strategy that coordinates scheduling, inventory, and production to move away from the batch mode of production in order to improve quality and reduce inventories. It is a paradigm shift which ensures that customers receive only what is needed, just when it is needed, and in the exact amounts needed.

**Kaizen:** continuous improvements in an activity over time which reduce waste. Such improvements are usually small and incremental in scope.

**Kaizen event:** a team event dedicated to quick implementation of a lean method in a particular area over a short time period.

**Kaizen plans:** lean improvement proposals presented to management by the team, following their analysis of their current state map and mapping their future state.

**Kanban:** a Japanese term which means “card signal”. However, this term is often used more generally to describe any visual signal use to indicate the need for material replenishment by an upstream activity.

**Kano model:** asserts that, for some customer requirements, satisfaction is proportional to the extent that the product or service is fully functional. Others are disproportional.

**Key performance indicators:** leverage variables contributing to business success. Entire set of KPI’s is also known as “system of indicators.”

**Knowledge transfer:** getting the right knowledge to the right people at the right time to achieve your purpose.

**Kuskal-Wallis test:** a nonparametric hypothesis test used to compare three or more independent samples.

**Latin squares:** a classical method used to generate orthogonal designs which are a subset of the full factorial. Typically used for screening strategies.



**LCL (lower control limit):** for control charts, the limit above which the process subgroup statistics ( $\bar{x}$ , R, etc.) must remain when the process is in control. Typically 3 standard deviations below the center line.

**Lead time:** the total time a customer must wait to receive a product after placing an order. When a scheduling and production system are running at or below capacity, lead time and throughput time are the same. When demand exceeds the capacity of a system, there is additional waiting time before the start of scheduling and production, and lead time exceeds throughput time.

**Lean:** body of knowledge and tools organizations use to remove all non-value-added time and activity from their processes. It is based on the Toyota Production System.

**Lean enterprise:** an organization that fully understands, communicates, implements, and sustains lean concepts seamlessly throughout all operational and functional areas.

**Lean production:** the principle philosophy of lean production is the elimination of waste, particularly as defined by the 8 types of waste. However, in a broader context, lean production refers to many of the techniques and methods which contribute to the company's ability to respond quickly and efficiently to customer demand by delivering high quality products when and where needed.

**Lean Six Sigma:** a business philosophy which combines the strategies of lean production (elimination of waste) and Six Sigma (reduction of variation). These concepts must be applied to all facets and functions of the business in order to deliver better products and services faster and at lower cost, while obtaining a maximum return on investment.

**Left-tailed test:** hypothesis test in which the critical region is located in the extreme left area of the probability distribution.

**Least squares:** a method of curve-fitting that defines the "best" fit as the one that minimizes the sum of the squared deviations of the data points from the fitted curve.

**Level:** a setting or testing value of a factor.

**Leveling:** evenly distributing over a shift or a day the work required to fulfill customer demand. Leveling is achieved either through implementing visible pitch or heijunka (load leveling).



**Level of significance:** a measure of the outcome of a hypothesis test. It is the P-value or probability of making a Type I error.

**Linear correlation:** the extent to which two variables correlate in a linear manner. That is, how close their scatterplot is to a straight line.

**Linear graph:** a tool used by Taguchi to identify sets of interacting columns in orthogonal arrays.

**Line balancing:** a process in which work elements are evenly distributed within a value stream to meet the takt time.

**Line chart:** charts use to track the performance without relationship to process capability or control limits.

**Little's law:** formula that shows the relationship between lead time, work in process (WIP), and completion rate.

**Load chart:** a vertical bar chart which compares the workload, in units of time, assigned to each operator or piece of equipment in a work cell. The cycle times within the cell should be balanced so that they are approximately equal, thereby minimizing idle time and delays. The load chart should reflect the overall process cycle time which equals the takt time.

**Location indicator:** a visual workplace element that shows where an item belongs. Lines, arrows, labels, and signboards are all examples of location indicators.

**Loss:** the reduction in value of a system, product, or process that does not function as it was intended.

**Loss function:** a technique for quantifying loss due to product deviations from target values.

**Lower confidence limit:** the smaller of the two numbers that form a confidence interval.

**Lower control limit:** a horizontal dotted line plotted on a control chart which represents the lower process limit capabilities of a process.

**Make a tradeoff:** a solution that finds a compromise between conflicting requirements.

**Matrix:** a table that describes the treatment combinations to be run. The usual convention displays the factors in columns and the treatment combinations in rows.



**Main effect:** the influence a single factor has on the response when it is changed from one level to another. Often used to represent the “linear effect” associated with a factor.

**Master Black Belt:** teaches Black Belts and Green Belts to become more proficient in executing Lean Six Sigma projects. Also coaches, mentors, and advises management.

**Maximum error of estimate:** the largest difference between a point estimate and the true value of a population parameter.

**Mean square error:** a weighted average of the variances for each run.

**Mean:** the average of a set of values.

**Measure of central tendency:** numerical measures that depict the center of a data set. The most commonly used measures are the mean and the median.

**Measure of dispersion:** any of several measures designed to reflect the amount of variability among a set of values.

**Measurement System Analysis (MSA):** a statistical technique that quantifies the amount of the variability that originates from the measurement system itself. In this context, the measurement system consists of the people and equipment needed to obtain the measurement, as well as all of the steps in the measurement process.

**Median:** the middle value of a data set when the values are arranged in either ascending or descending order.

**Midquartile:** one-half of the sum of the first and third quartiles.

**Midrange:** one-half the sum of the highest and lowest scores.

**Milestones:** an activity that has, or should have, a scheduled completion date or time, and that has a major effect on the timing or completion of a project.

**Mindmapping:** brainstorming technique for diagramming concepts and relationships that stem from an original problem statement.

**Mistake proofing:** “Poke-yoke.” Term coined by Shigeo Shingo to describe the technique of changing a process to prevent mistakes from occurring.

**Mixed effects model:** contains elements of both the fixed and random effects models.



**Mixture design:** a special type of study in which the factors are ingredients that are mixed together. The response variables are thought to depend on the relative proportions of the components of the mixture rather than the absolute concentrations.

**Mode:** the value which has occurred the highest number of times in a set of data.

**Monte Carlo Simulation:** uses mathematical models, probability, decision rules, and induced variability in inputs to simulate real-life scenarios and identify variability in outputs.

**Monument:** any design, scheduling, or production technology with scale requirements necessitating that designs, order, and products be brought to the machine to wait in a queue for processing.

**MTBF (mean time between failures)** mean time between successive failures of a repairable product. This is a measure of product reliability.

**Muda:** any activity which is wasteful. Muda uses time or resources, but does not add any value from the perspective of the customer.

**Multicollinearity:** the existence of strong correlations between input factors or independent variables.

**Multimodal:** having more than two modes.

**Multinomial experiment:** an experiment with a fixed number of independent trials and each outcome falls into exactly one of several categories.

**Multiple coefficient of determination:** measure of how well a multiple regression equation fits the sample data.

**Multiple correlation:** the correlation of one variable with a combination of other variables.

**Multiple regression:** a model where several independent variables are used to predict one dependent variable.

**Multivariate analysis:** statistical analysis involving more variables than one independent variable and one dependent variable.

**Multi-vari diagram:** a graphical means of showing stratified data.

**Multivoting:** a procedure for prioritizing problems, ideas, etc., whereby votes are taken in stages among the stakeholders, number of items being reduced at each stage.

**Mura:** unevenness in operations (poor flow) contributing to waste (muda).

**Muri:** overburden of people or machines contributing to waste (muda).

**Mutually exclusive events:** events that cannot both happen simultaneously.

**Natural tolerances (of a process):** 3 standard deviations on either side of the center point (mean value). In a normally distributed process, the natural tolerances encompass 99.73% of all measurements.

**Nested or hierarchal design:** an experiment to examine the effect of two or more factors in which the same level of a factor cannot be used with all levels of other factors.

**Noise factors:** factors that engineers cannot control on a regular basis, or prefer not to control. Noises (i.e., ambient temperature) cause a product's response to vary from the target performance. Expensive factors are considered to be noise during parameter design.

**Nominal:** for a product whose size is of concern: the desired mean value for the particular dimension, the target value. Also a level of measurement of data in which the data consist of names, labels, or categories only.

**Nominal data:** when we use numbers as labels for categories we refer to the data collected as nominal. We cannot perform mathematical operations on these numbers. For example, if we label the category 'men' as 1 and 'women' as 2, we cannot add up two men and claim it equals one woman! The data is usually the number of subjects in each category.

**Non-additive:** the total effect of several factors does not equal (exactly) the sum of the individual factor effects. Interactions imply non-additivity.

**Nonconforming unit:** a sample (part) which has one or more nonconformities, making the sample unacceptable for its intended use.

**Nonlinear effect:** a change in a response variable that is not linearly related to the corresponding change in a factor.

**Nonparametric test:** statistical tests that do not use, or make assumptions about, the characteristics (parameters) of populations.

**Non-Value Added (NVA):** activities which consume time or resources but do not directly contribute towards meeting the customer's requirements.



**Normal distribution:** the distribution characterized by the smooth, bell-shaped curve.

**n-p chart:** for attribute data: a control chart of the number of defective units in a subgroup. Assumes a constant subgroup size. Based on the binomial distribution.

**Nuisance variable:** an unknown background variable that can affect a response variable in an experiment, sometimes called a lurking variable or extraneous variable.

**Null hypothesis ( $H_0$ ):** the conclusion that typically includes equality, i.e.,  $H_0: \mu_1 = \mu_2$  or  $H_0: \sigma_1 = \sigma_2$ .

**Observed frequency:** the actual frequency count recorded in one cell of a contingency table or multinomial table.

**One-tailed test:** a prediction that two samples come from different populations, specifying the direction of the difference, that is, which of the two populations will have the larger mean value.

**One-way analysis of variance:** analysis of variance involving data classified into groups according to a single criterion only.

**Operational definition:** a definition that gives communicable meaning to a concept by specifying how the concept is applied within a particular set of circumstances.

**Operation:** any activity or activities performed on a product by a single machine.

**Optimization:** the employment of methodologies to improve the quality, reduce cost, and shorten development time of products and processes. Optimization is a balance between the voice of the customer and the voice of the business.

**Ordinal data:** when we cannot assume that the intervals between consecutive numbers on a scale of measurements are of equal size we have ordinal data and can only use the data to rank order the subjects. Ratings are assumed to be ordinal data. We perform nonparametric tests on ordinal data.

**Orthogonal:** a design is orthogonal if the main and interaction effects in a given design can be estimated without confounding the other main effects or interactions. A full factorial is said to be balanced, or orthogonal, because there are an equal number of data points under each level of each factor.

**Outcome:** a possible result from an experiment. It may also be referred to as a basic outcome.



**Outer array:** in a Taguchi style fractional factorial experiment, these are the factors that cannot be controlled in a process.

**Outlier:** a data point which is not part of the natural clustering of the rest of the data set.

**Output:** the product, service, or task resulting from the completion of a process. It also is used to denote a measure of performance of a product, service, or task.

**Out of control (of a process):** a process is said to be out of control if it exhibits variations larger than its control limits, or shows a systematic pattern of variation.

**Overall Equipment Effectiveness (OEE):** a measure of equipment performance in terms of downtime, throughput relative to design specifications, and the quality of the output.

**p-chart (percent defective):** for attribute data: a control chart of the proportion of defective units (or fraction defective) in a subgroup. Based on the binomial distribution.

**P-value:** the probability of making a Type I error. This value comes from the data itself. It also provides the exact level of significance of a hypothesis test.

**Paired comparison experiments:** a single factor at two levels and on background variable.

**Parameter:** a numerical measure of some aspect of a population.

**Parameter design:** a strategy intended to reduce the effect of all noise variables on quality characteristics for a product or process.

**Parametric test:** statistical tests that use the characteristics (parameters) of populations or estimates of them.

**Pareto diagram:** a bar chart for attribute (or categorical) data that is presented in descending order of frequency.

**Percent defective:** for acceptance sampling: the percentage of units in a lot which are defective, i.e., of unacceptable quality.

**Percentiles:** a method used to describe the location of values within a data set. It divides the entire range of a data set into 100 equal parts called percentiles.

**Performance dashboard:** group of hierarchically linked performance metrics monitored on a regular basis.





**Performance index:** any measure that indicates the health of a business.

**Permutations rule:** rule for determining the number of different arrangements of selected items.

**PF/CE/CNX/SOP:** a powerful methodology that can be used to remove up to 70 – 80% of extraneous process variation. It employs process flow (PF) diagrams with cause and effect (CE) diagrams to identify and sort the causes of variation. The causes are further categorized as controllable (“C”), noise (“N”) or experimental (“X”) variables. Standard operating procedures (SOPs) are used to control a “C” variable and hold it as constant as possible.

**Physical process map:** a layout diagram of the work area that illustrates the path followed by materials or parts through the facility. It can be used to highlight excessive amounts of material transport or employee movement while following the steps described in a process flow chart. Synonymous with spaghetti diagram.

**Pictograph:** a tool that uses a sketch of a part, a plant, a city, etc. to mark the exact location of problem occurrences.

**Pie chart:** graphical method for representing data in the form of a circle containing wedges.

**Pilot:** small scale test of a proposed change applied to the real process, not a hypothetical or simulated process.

**Pitch:** a multiple of takt time that will allow you to create, maintain, and sustain a consistent and practical workflow throughout the value stream. To calculate pitch, multiply the takt time by the number of work units to flow through the system in a manageable way.

**Planned experimentation:** a collection of methods and a strategy to make a change to a product or process and observe the effect of that change on one or more quality characteristics with the purpose of helping experimenters gain the most information with the resources available.

**Planned grouping:** arrangement of experimental units in blocks.

**Point estimate:** the use of a sample statistic (or single value) such as  $\bar{x}$  to estimate a population parameter  $\mu$ .

**Poisson distribution:** a probability distribution for the number of occurrences per unit interval.



**Poka-yoke:** any technique(s) which is used to reduce the likelihood of an error through color coding, unique design characteristics which prevent improper assembly, etc. These techniques are best applied in the design phase of a process to prevent errors but can also be implemented as a corrective action to prevent recurrences of an error.

**Population:** a set or collection of objects or individuals. It can also be the corresponding set of values which measure a certain characteristic of a set of objects or individuals.

**Power of a test:** the probability that, when there is a genuine effect to be found, the test will find it (that is correctly reject a false null hypothesis). As an illustration, one test might be a stopwatch that gives the same time for two runners in a race but a more powerful test is a sensitive electronic timer that more accurately shows the times to differ by a fiftieth of a second.

**Precision of a measurement process:** the degree of variation in individual measurements of the same item.

**Pre-control:** a method of controlling a process based on the specification limits. It is used to prevent the manufacture of defective units, but does not work toward minimizing variation of the process. The area between the specifications are split into zones (green, yellow and red) and adjustments made when a specified number of points fall in the yellow or red zones.

**Predicted value:** using a regression equation, the value of one variable given a value for the other variable.

**Prediction:** a declaration of the value or state of some characteristic of a process or its outcome in the future.

**Preliminary design:** concept description sufficiently detailed to permit optimization and validation.

**Prevention:** the practice of eliminating unwanted variation a priori (before the fact), e.g., predicting a future condition from a control chart and then applying corrective action before the predicted event transpires.

**Primary control variables:** the major independent variables used in the experiment.

**Prioritization matrix:** a tool for comparing problems or solutions by weighing each against a set of criteria.

**Probability:** the chance of a specific event occurring from a set of possible events, expressed as a proportion. For example, if there were 4 women and 6 men in a room the probability of meeting a woman first on entering the room is 4/10 or 0.4 as there are 4



women out of 10 in the room. A probability of 0 indicates an even will never occur and a probability of 1 that it will always occur.

**Probability distribution:** a table, graph, or formula that describes the probabilities of all of the possible outcomes of an experiment. Many naturally occurring phenomena and processes can be described by a probability distribution. Knowing the distributional form that is associated with a specific experiment, process, or phenomena allows us to make accurate prediction regarding the likelihood of different events related to that experiment, process, or phenomena.

**Problem:** a deviation from a specified standard.

**Problem linkage:** the notion that some identified problems might be the cause or the effect of other identified problems.

**Problem solving:** a team working together to follow these steps – defining the problem; measuring the problem; analyzing possible causes; identifying possible solutions; developing an action plan; evaluating and renewing the action plan; standardizing effective ideas.

**Process:** a series of activities, administrative or production-oriented, which are usually performed in a particular sequence to accomplish a specific objective or produce a desired product.

**Process average:** the central tendency of a given process characteristic across a given amount of time or at a specific point in time.

**Process capability:** a measure of the quality of a process (and its output) that is derived through a comparison of the process variation to the process specification limits. Capability measures include sigma level, sigma capability,  $C_p$ ,  $C_{pk}$  and defects per million (dpm).

**Process control chart:** a fundamental tool of statistical process control (SPC) which is used to determine if a process is statistically under control. This implies stability and predictability within the bounds of natural variation. It does not imply process capability.

**Process flow chart:** a diagram that describes in sequence all of the major steps and decision points in a process. It improves understanding of the process and can facilitate the identification of problem areas and sources of variation.

**Process improvement:** the continuous endeavor to learn about the cause system in a process and to use this knowledge to change the process to reduce variation and complexity and to improve customer satisfaction.



**Process observation chart:** a table used to record specific attributes of each step while carefully watching the process. The attributes of interest usually include the type of process step, including operation, transportation, inspection, waiting, storage or decision, quantity, distance, and elapsed time.

**Process owner:** the person at the lowest level in the hierarchy of the organization who has the authority to make fundamental changes to the process.

**Process spread:** the range of values which a given process characteristic displays; this particular term most often applies to the range but may also encompass the variance. The spread may be based on a set of data collected at a specific point in time or may reflect the variability across a given amount of time.

**Processing time:** the time a product is actually being worked on in design or production and the time an order is actually being processed. Typically, processing time is a small fraction of throughput time and lead time.

**Producers risk:** the probability of rejecting a lot when, in fact, the lot should have been accepted (see alpha risk).

**Productivity:** doing more with the same or fewer resources.

**Project charter:** the guiding document for a Lean Six Sigma project.

**Product requirements:** the suite of customers' requirements, regulatory requirements, marketing requirements, industry standards and internal standards which specify cost, delivery, operability, environmental resistance and durability of a product.

**Pugh analysis:** a tool for comparing the merits and demerits of several potential solutions.

**Pull:** a strategy which requires downstream process steps to signal a need for replenishment, before activity in each of the previous (upstream) process steps is initiated. In the ideal pull system, no production begins until the customer purchases the product, and even then, it is only produced in the quantity required. The customer demand "pulls" the product through the process.

**Push:** conventional work in which schedules are pushed along based on sales projections and availability of materials. It leads employees to make as many work units as they can as fast as they can, even if the next process is not ready to use the work units, which causes long queue times.

**p-value:** the probability that a test statistic in a hypothesis test is at least as extreme as the one actually obtained.



**Qualitative:** this refers to descriptors of category and/or order, but not of interval or origin. Different machines, operators, materials, etc. represent qualitative levels or treatments.

**Quality characteristics:** a trait, preferably measurable, of an input or outcome of a process or a measure of performance of a process used to define quality.

**Quality Function Deployment (QFD):** a systematic process used to translate and integrate the “voice of the customer” into the development of products or services that will meet the customer requirements.

**Quality rate:** the percentage of time you meet performance standards the first time. Also known as “first-time-yield” (FTY).

**Quantitative:** this refers to descriptors of order and interval (interval scale) and possibly also of origin (ratio scale). As a quantitative factor, temperature might describe the interval value 27.32 deg. C. As a quantitative response, yield might describe the ratio value 87.42%.

**Quartile:** if we order a set of values from the lowest to the highest the quartiles are the points that divide the values into four equal groups, with a quarter of the values in each group. The second quartile is the median.

**Queue time:** the time a product spends in a line awaiting the next design, order-processing, or fabrication step.

**R charts:** plot of the difference between the highest and lowest values in a sample. Range control chart.

**Random:** selecting a sample so each term in the population has an equal chance of being selected; lack of predictability; without pattern.

**Random cause:** a source of variation which is random; a change in the source (“trivial many” variables) will not produce a highly predictable change in the response (dependent variable), e.g., a correlation does not exist; any individual source of variation results in a small amount of variation in the response; cannot be economically eliminated from a process; an inherent natural source of variation.

**Random effects model:** experimental treatments are a random sample from a larger population of treatments. Conclusions can be extended to the population. Interferences are not restricted to the experimental levels.



**Random error:** there will always be random factors influencing subjects' values in an experiment. Random error is the influence of these random factors on the data. Statistical tests take account of random factors.

**Random order:** randomization of treatment combinations.

**Random sample:** a sample of a population where each member of the population has an equal chance of being chosen for the sample.

**Random variable:** a definition of the possible outcomes of interest from a given experiment.

**Random variations:** variations in data which result from causes which cannot be pinpointed or controlled.

**Random word:** a creative thinking tool in which one uses a random word out of a dictionary as a thought provoker to generate new thought patterns.

**Randomization:** the objective assignment of combination of factor levels to experimental units.

**Randomized block design:** a experimental pattern in which the size of the block equals the number of combinations of factor and levels (either a single background variable or a chunk variable).

**Randomized trials:** frees an experiment from the environment and eliminates biases. This technique avoids the undue influences of systematic changes that are known or unknown.

**Range:** the difference between the highest and lowest value in a set of data.

**Rank:** when a set of data is ordered from lowest to highest the rank of a value is its position in this order.

**Rank correlation coefficient:** measure of the strength of the relationship between two variables; based on the ranks of the values.

**Rank order:** a method of ordering values, listing them from lowest to highest.

**Rapid changeover:** lean practice of quickly changing over from one mode of production (or service) to another. Also called "Single Minute Exchange of Die" (SMED).

**Ratio data:** data measured on a ratio scale.



**Ratio scale:** an interval scale with an absolute zero. A stopwatch has an absolute zero a 0 indicates 'no time' and so we can make ration statements: 20 seconds is twice as long as 10 seconds.

**Red-tag event:** the process of removing unneeded items in order to complete the sorting step of the Five S's.

**Reduce the effect:** to make a system more robust to detrimental causes or sources of problems.

**Regression:** the prediction of values on one variable by their values on a second variable. The larger the correlation between the variables the more accurate the prediction.

**Regression line:** a straight line that summarizes the relationship between two variables.

**Reject region:** the region of values for which the alternate hypothesis is accepted.

**Relations diagram:** an arrow diagram showing the cause and effect relationships among problems, performance indices, etc.

**Reliability apportionment tree:** a tree diagram that shows how reliable each component of an assembly must be in order to meet the overall reliability target.

**Repeatability:** (precision) the variation between successive measurements of the same part, same characteristic, by the same person using the same instrument. Also, known as test-retest error.

**Repeated trails:** trials that are conducted to estimate the pure trial-to-trial experimental error so that lack of fit may be judged. Also called replications.

**Repetition:** repeated observations of the same treatment combination without resetting the factor levels.

**Replicate:** a repeat run of treatment combination. The factor levels are completely reset for each factor level.

**Replication:** observations made under identical test conditions.

**Representative sample:** a subset of a population that shares the same key characteristics of the population. For example, the sample has the same ration of men to women as the population.





**Reproducibility:** the difference in the average of the measurements made by different persons using the same or different instrument when measuring the identical characteristic.

**Requirements matrix:** a tool used in QFD to show the relationships between customer requirements and technical characteristics.

**Residual:** a linear regression provides a prediction of the subjects' values on one variable by their values on a second. The residual is the difference between the subjects' actual value and their predicted value on the first variable.

**Residual error:** the difference between the observed and the predicted value for that result, based on an empirically determined model. It can be variation in outcomes of virtually identical test conditions.

**Resolution I:** an experiment in which tests are conducted, adjusting one factor at a time, hoping for the best. This experiment is not statistically sound.

**Resolution II:** an experiment in which some of the main effects are confounded. This is very undesirable.

**Resolution III:** a fractional factorial design in which no main effects are confounded with each other but the main effects and two factor interactions effects are confounded.

**Resolution IV:** a fractional factorial design in which the main effects and two factor interaction effects are not confounded but the two factor effects may be confounded with each other.

**Resolution V:** a fractional factorial design in which no confounding of main effects and two factor interactions occurs. However, two factor interactions may be confounded with three factor and higher interactions.

**Resolution VI:** also called resolution V+. This is at least a full factorial experiment with no confounding. It can also mean two blocks of 16 runs.

**Resolution VII:** can refer to eight blocks of 8 runs.

**Response:** what you measure; the quality characteristic of the product. The response is also called the dependent variable.

**Response plots:** a plot illustrating the relationship between the response and important factors in an experiment.



**Response surface methodology:** a tool for analyzing how an output variable will respond to several input variables; it uses design of experiments to collect data.

**Response variable:** a variable observed or measured in an experiment, sometimes called a dependent variable. The response variable is the outcome of an experiment and is often a quality characteristic or a measure of performance of the process.

**Right-tailed test:** hypothesis test in which the critical region is located in the extreme right area of the probability distribution.

**Risk priority number (RPN):** in a FMEA, the product of a potential failure mode's Severity, Occurrence, and Detection ratings.

**Robust:** the condition or state in which a response parameter exhibits hermeticity to external cause of a nonrandom nature; i.e., insensitive to noise.

**Robust design:** a term associated with the application of Taguchi experimentation in which a response variable is considered robust or immune to input variables that may be difficult or impossible to control.

**Rolled throughput yield:** the probability that a part will make it through multiple process steps without a defect.

**Root cause:** the x (cause) that possesses the most leverage in determining the state of the desired outcome (effect).

**Run:** a test of a treatment combination.

**Run chart:** a graphical tool that charts a process performance measured over time.

**Run order:** order in which the treatment combinations are tested.

**Run test:** nonparametric method used to test for randomness.

**Runner:** a worker who ensures that pitch is maintained. The runner covers a designated route within the pitch period, picking up work units, folders, or kanban cards, and delivering them to their appropriate places.

**Sample:** a set of objects, individuals or values selected from the population.

**Sample space:** in an experiment, the set of all possible outcomes or events that cannot be further broken down.

**Sampling:** the process of obtaining parts or observations for analysis.

**Sampling errors:** errors resulting from the sampling process itself.

**Safety resources:** a means of meeting customer demand when internal constraints or inefficiencies disrupt process flow.

**Scatterplot:** a graph of values on one variable plotted against their values on a second variable. The graph shows how the values are ‘scattered.’

**Screening designs:** a set of fractional factorial designs used by the experimenter with a low level of knowledge to screen out unimportant factors.

**Separation:** a design principle in which conflicting requirements can be met by physically separating the elements or components of the design.

**Sequential experimentation:** sequential building of knowledge using the improvement cycle (PDSA) with prediction as the aim.

**Set in order:** the second activity in the 5S system. It involves identifying the best location for each item that remains in the area, relocating items that do not belong in the area, setting height and size limits, and installing temporary location indicators.

**Shewhart cycle:** the cycle of Plan/Do/Study/Act developed by statistician Walter Shewhart. Also known as the Plan/Do/Check/Act cycle.

**Shift:** a dramatic change in the mean of a process metric.

**Shift and drift:** tendency for process variation to shift 1.5 sigma from the short to the long run. Also called the 1.5 sigma shift.

**Shift and squeeze:** a process improvement strategy to shift the mean to the desired level and to reduce its associated variation.

**Shine:** The third activity in the 5S system. It involves cleaning everything thoroughly, using cleaning as a form of inspection, and coming up with ways to prevent dirt, grime, and other contaminants from accumulating.

**Sigma:** a Greek letter  $\sigma$  used to represent the standard deviation of the population. It is a measure of the variability in the data.

**Sigma level:** a commonly used measure of process capability that represents the number of standard deviations between the process center and the closest specification limit.



**Signal to noise ratio:** the signal to noise ration (S/N) is one method of incorporating the loss function into experimental design for robust design.

**Significance level:** the risk (probability) of erroneously claiming a relationship between an independent and dependent variable when there is not one. Statistical tests are undertaken so that this probability is chosen to be small, usually set at 0.05 indicating that this will only occur no more than 5 times in 100.

**Significance test:** any test, such as a t-test, for determining through data whether one condition is different from another.

**Sign test:** a nonparametric hypothesis test used to compare samples from two populations.

**Simple linear regression:** a mathematical model where one independent variable is used to predict the value of a dependent variable using a linear equation (i.e., no quadratics).

**Simple main effects:** a significant interaction in a two factor analysis of variance indicates that the effect of one variable is different at the various conditions of the other variable. Calculating simple main effects tell us what these different effects are. A simple main effect is the effect of one variable at a single condition of the other variable.

**Simple solution:** a design change that has no detrimental effects.

**Simplex design:** a spatial design used to determine the most desirable variable combination (proportions) in a mixture.

**Simulation:** experimentation used to manipulate critical x's to see how they impact the overall process.

**Single Minute Exchange of Dies (SMED):** a series of techniques used to facilitate the rapid changeover of any process or equipment.

**Single piece flow:** a production strategy in which individual units progress as a batch of one unit through each of its associated process steps with continuous flow in the forward direction and without rework. It can be contrasted with the batch and queue strategy in which batches of multiple units proceed as a group through each of the production steps. This procedure requires each unit to stop and wait for each of the other units in the batch to be completed before the group proceeds to the next process step.

**SIPOC:** a tool for identifying the scope of a project. It stands for Suppliers, Inputs, Process, Outputs, and Customers.



**Six Sigma:** a quality improvement and business strategy that began in the 1980's at Motorola. Emphasis is on reducing defects, reducing cycle time with aggressive goals, and reducing costs to dramatically impact the bottom line.

**Slope:** measure of steepness of a straight line.

**Soft savings:** future savings, or avoided costs (as opposed to current-year "hard" savings); or, intangible benefits like "employee satisfaction."

**Sort:** The first activity in the 5S system. It involves sorting through and sorting out items, placing red tags on these items, and moving them to a temporary holding area. The items are disposed of, sold, moved, or given away by a predetermined time.

**Spaghetti chart:** a layout diagram of the work area that illustrates the path followed by materials or parts through the facility. It can be used to highlight excessive amounts of material transport or employee movement while following the steps described in a process flow chart. Synonymous with physical process map.

**Spearman's rank correlation coefficient:** see rank correlation coefficient.

**Special cause variation:** non-random causes of variation that can be detected by the use of control charts and good process documentation. A process is said to be in a state of statistical control when all sources of special cause variation have been eliminated.

**Specifications:** exacting requirements for CT's (Critical-To characteristics) expressed mathematically so they can be constantly measured and monitored.

**Stability:** the absence of assignable causes of variation; the property of being in statistical control. The consistency of measurements over time.

**Stable process:** a process which is free of assignable causes, e.g., in statistical control.

**Standard deviation:** a measure of the standard ('average') difference of a value from the mean in a set of data. It is the square root of the variance.

**Standard error of the estimate:** a measure of the 'average' distance (standard error) of a value from the regression line.

**Standard error of the mean:** the standard deviation of the distribution of sample means. It is a measure of the standard ('average') difference of a sample mean from the mean of all samples of the same size from the same population.



**Standardize:** the fourth activity in the 5S system. It involves creating the rules for maintaining and controlling the conditions established after implementing the first three S's. Visual controls are used to make these conditions obvious.

**Standardized work:** an agreed-upon set of work procedures that establishes the best method and sequence for each process. Standardized work is implemented to maximize efficiency while simultaneously ensuring safe conditions.

**Standard normal distribution:** a normal distribution with a mean of 0 and standard deviation of 1.

**Standard Operating Procedure (SOP):** an up-to-date written procedure that clearly and concisely describes the exact method to be followed in order to complete a specific task.

**Standard score:** the position of a value within a distribution of values. It provides a measure of how many standard deviations a specific value falls above or below the mean. It is also referred to a z-score.

**Standard work:** a method of improving work efficiency and reducing variability by insuring that everyone follows exactly the same procedure to complete a specific task. The procedure may be improved as new ideas or technology becomes available. However, the modified procedure must first be demonstrated to be superior and if adopted as the best practice, must be followed by everyone performing that task.

**Statistic:** specifically, a characteristic of a sample, such as the sample mean. More generally, statistic and statistics are use to describe techniques for summarizing and analyzing numerical data.

**Statistical control:** a quantitative condition which describes a process that is free of assignable/special causes of variation, e.g., variation in the central tendency and variance. Such a condition is most often evidenced on a control chart, i.e., a control chart which displays an absence of nonrandom variation.

**Statistical process control:** the use of graphical and statistical methods to analyze and improve a process by reducing its variation and increasing its capability.

**Statistics:** the collection, organization, description, and analysis of data.

**Stem-and-leaf plot:** method of sorting and arranging data to reveal the distribution.

**Stepping stone:** a creative thinking tool in which one uses a less-than-perfect brainstorming idea as a thought provoker for new ideas.



**Stepwise regression:** in multiple regression, the process of using different combinations of variables until the best model is obtained.

**Stratified sampling:** samples are drawn from each stratum or class.

**Storyboard:** a poster-sized framework for holding all the key information for a lean implementation. It contains the outcomes for each of the eight steps of value stream management.

**Subgroup:** a sample of units drawn from a process at, or near, the same time. Data collected from measurements of the units are used to plot control charts.

**Sum of squares:** the sum of the squared deviations of values from their mean value.

**Supermarket:** a system used to store a set level of finished-goods inventory or WIP and replenish what is “pulled” to fulfill customer orders (internal and external). A supermarket is used when circumstances make it difficult to sustain continuous flow.

**Supplier:** person or organization that provides inputs to a process.

**Sustain:** the fifth activity of the 5S system, where a person or team ensures adherence to 5S standards through communication, training, and self-discipline.

**Swim lane flow chart:** displays the steps of a process and shows functional ownership for each one.

**Symptom:** that which serves as evidence of something not seen.

**Systematic error:** data that has been systematically influenced by another variable in addition to the independent variable under test is said to contain systematic error. The additional variable is said to confound the experiment.

**Systematic sampling:** every kth element is selected for a sample.

**Synthesis:** concerned with how it works, a study of the whole (system) to understand the parts. Commonly associated with parameter design.

**t-distribution:** a bell-shaped distribution usually associated with small sample experiments. Also called the Student t distribution.

**T-test:** a statistical test that shows whether samples from condition A are different from samples from condition B.





**Take for granted analysis:** a creative thinking tool in which one lists everything that is taken for granted concerning a concept, and then analyzes the validity of these assumptions.

**Takt time:** the average time to produce a unit in order to meet the current rate of customer demand. Takt time determines how fast a process needs to run to meet customer demand. Takt time is calculated by dividing the total operating time available by the total quantity required by the customer.

**Team charter:** a document that includes but is not limited to the following elements: 1) a clear definition of a team's mission, 2) a statement of team members' roles and responsibilities, 3) a description of the scope of the team's responsibilities and authority, 4) project deadlines, 5) a list of metrics and targets, and 6) a list of deliverables.

**Technical characteristics:** an objective measure of a quality requirement.

**Test of significance:** a procedure to determine whether a quantity subjected to random variation differs from a postulated value by an amount greater than that due to random variation alone.

**Test statistic:** used in hypothesis testing, it is the sample statistic based on the sample data.

**Tests for root cause:** there are three: (1) Is the root cause in the necessary state to cause the problem? (2) Is there no other deeper root cause? (3) Can the problem be turned on and off via the root cause?

**Theory of Inventive Problem Solving (TRIZ):** an innovative methodology that uses a proven matrix of generic solutions to solve specific problems.

**Thought process map:** a tool for identifying and asking the questions necessary to complete a project.

**Throughput time:** the time required for a product to proceed from concept to launch, order to delivery, or raw materials into the hands of the customer. This includes both processing and queue time.

**Time value map:** a time scaled graph that illustrates the amount of active and inactive time elapsed during one complete cycle of a process.

**Tolerance design:** a strategy that helps set acceptable levels of variation for control factors (considered costs) of a product or process after parameter design has been performed.





**Tollgating:** process for regularly reviewing Lean Six Sigma projects at the end of each successive phase of DMAIC. Also called tollgate review.

**Top-down flowchart:** a diagram depicting process flow horizontally, with detailed instructions listed below each.

**Total cycle time:** the total of the cycle times for each individual operation or process or work area in a value stream. Total cycle time ideally equals total value-added time.

**Total lead time:** the total of all cycle times from all individual processes within the lean value stream, plus the queue times that exist between each process.

**Total deviation:** the sum of the explained deviation and unexplained deviation for a given pair of values in a collection of bivariate data.

**Total Productive Maintenance (TPM):** a comprehensive and coordinated maintenance program designed to maximize equipment effectiveness by minimizing downtime and optimizing output in terms of speed and quality. This approach is founded upon the ability of well trained equipment operators to proactively identify and correct small maintenance issues before they become significant and lead to breakdowns.

**Total variation:** the sum of the squares of the total deviation for all pairs of bivariate data in a sample.

**Transfer function:** the mathematical relationship between a set of process inputs and a specific output feature.

**Treatment combination:** refers to how the factors will be set for the experimental run.

**Tree diagram:** a tool that is structured like a tree or a tree's roots, that shows the familiar relationships between problems, ideas, components, etc.

**Tuning:** tuning is the process of setting the tuning factors at the proper settings to meet a set of requirements. Tuning is used to adjust a generic product to a specific application.

**Tuning research:** tuning research is the process of identifying the factors which affect the product/process output and finding their practical limits while remaining robust.

**Two-tailed test:** a prediction that two samples come from different populations but does not state which population has the higher mean value.

**Two-way analysis of variance:** analysis of variance involving data classified according to two different factors.



**Type I error:** the error of accepting the null hypothesis when it is true. The risk of this occurring is set by the significance level.

**Type II error:** the error of accepting the null hypothesis when it is false.

**Unexplained deviation:** for one pair of values in a collection of bivariate data: the difference between y coordinate and predicted value.

**Unexplained variation:** the sum of squares of the unexplained deviations for all pairs of bivariate data in a sample.

**Uniform distribution:** a distribution of values evenly distributed over the range of possibilities.

**Unit-to-unit noise factors:** variations in the manufacturing process.

**Universe:** the entire group (e.g., people, material, invoices) possessing certain properties of interest.

**Unnatural pattern:** any pattern in which a significant number of the measurements do not group themselves around a center line; when the pattern is unnatural, it means that outside disturbances are present and are affecting the process.

**Unstable process:** a process in which variation is a result of both common and special causes.

**Upper control limit:** a horizontal line on a control chart (usually dotted) which represents the upper limits of process capability.

**Validation:** the process which produces documented evidence that the product/process is capable of consistently yielding deliverables which meet customer requirements.

**Value:** a feature, condition, service, or product that the customer considers desirable and which is delivered to them when and where they want it. Value is measured in terms of what an external customer is willing to pay for.

**Value added percentage:** the percentage of the total lead time that is spent actually adding value to a work unit. To calculate value added percentage, divide the total cycle time by the total lead time.

**Value stream:** all of the steps, tasks and activities which are required to progress a product from receipt of raw materials to finished goods inventory, from receipt of a customer order to delivery of that product or service to the customer, or from research, development, engineering through early rate pilot production.



**Value stream champion:** the person with the authority and responsibility to allocate the organization's resources during the life of the project. The champion should always be completely committed to the project. It is often the champion who initiates the project.

**Value stream management:** a sequential, eight step process used to implement lean concepts and tools derived from the Toyota Production System. The purpose of value stream management is to minimize the waste that prevents a smooth, continuous flow of product throughout the value stream.

**Value stream mapping:** a diagram which describes all of the activities in a product or process value stream.

**Variables:** metrics (e.g., performance measures, inputs, outputs) whose values are subject to change or variability.

**Variables data:** a measurement on a continuous scale whose value is only limited by the sensitivity and resolution of the measuring system. It is the opposite of attribute data.

**Variance:** a measure of how much a set of values vary from their mean value. Variance is the square of the standard deviation.

**Variance components:** estimates of the variation due to each factor in a nested study.

**Variation:** any quantifiable difference between individual measurements; such differences can be classified as being due to common causes (random) or special causes (assignable).

**Visible pitch board:** a visual control method that will control the flow of work throughout the day, also showing how each person in the value stream receives his or her portion of the work.

**Visual management:** a series of visual techniques which are used to communicate the status of a system in such a way that it can be understood at a glance by everyone concerned. The term "system" is used broadly to describe everything from simple conditions such as the presence or absence of the tools required to perform a specific task such as a changeover system, or more complex operations such as graphs reporting key performance indicators for a business.

**Voice of the customer:** a term which describes the information, as well as collecting the information, about what is important to the customer. It includes but is not limited to: customer specifications, customer design requirements, customer surveys, and listening to customer feedback.



**Waste:** any activity or process that does not add value to the product or service from the perspective of the customer.

**Waste analysis:** an examination of a process that is conducted to identify and separate value added activities from non-value added activities or waste.

**Wilcoxon rank-sum test:** a nonparametric hypothesis test used to compare two independent samples.

**Wilcoxon signed-ranks test:** a nonparametric hypothesis test use to compare two dependent samples.

**Work cell:** a work area, often including machines, which has been designed to facilitate production activities by one or more operators. The physical layout and activities in a cell are very organized and carefully sequenced to optimize flow of materials with minimal effort.

**Work in Process (WIP):** material which has been partially processed, but has not yet reached the state of approved finished product.

**Work unit:** a specific, measurable amount of work that can be customized and treated as a whole. Examples of a work unit include an order, a report, an item, or a blueprint.

**Work unit family:** a group of parts that share common equipment and processing attributes.

**Worker balance chart:** a visual display of the work elements, times, and workers at each location. It is used to show improvement opportunities by visually displaying the times of each work operation in relation to the total value stream cycle time and takt time.

**X-bar and r charts:** a control chart which is a representation of process capability over time; displays the variability in the process average and range across time.

**Yellow Belt:** a less informed and trained Belt who collects data, characterizes process performance, and assists Black and Green Belts with project execution.

**Yield:** the proportion of “good” units produced relative to the theoretical maximum number of units which could be produced on the basis of material input quantities. Yield is usually expressed as a percentage of a theoretical maximum. It is also sometimes described in terms of absolute quantities such a kilograms of product produced, etc.

**y-intercept:** point at which a straight line crosses the y-axis.



**Z-distribution:** a standardized normal distribution or random variable having a mean of zero and a standard deviation of one.

**z score:** see standard score.

**Z-value:** a standardized value calculated by subtracting the mean from the recorded value and then dividing this difference by the standard deviation. It represents the number of standard deviations between a specified value and the mean.