Notes:

Data Collection Plan

Key Learning Points

1. Describe the importance of a data collection plan.
2. Explain how to develop a data collection plan.
3. Utilize data collection plans in improvement projects.

What is a Data Collection Plan?

Data and facts are fundamental to determine root causes of every improvement project. Without facts, problem-solving efforts are reduced to a "guessing game" with relatively low odds of success. A data collection plan is a tool used to define a clear strategy to efficiently collect reliable information that will be used to prove root causes.

- Data vs. Information: The team needs to understand the difference between data and information.
- Data Collection Cycle:The team needs to understand the model for generating useful information as shown.

Points of Use:

- Data collection is used at two steps during the DMAIC process:
- When measuring baseline performance of the Y in the Measure Phase.
- When collecting data in Analyze to Prove Xs.

Example Data Collection Plan

Healthcare

What to Measure		Operational Definition		Data Collection	
Question to Answer	Measure Name	Words	Formulas	Data Type	Sample Size
What percentage of patients are financially cleared by day of service?	Financial Clearance (Y)	Percentage of patients with all necessary and sufficient financial clearance elements completed by day of service	(# Cleared / Total) x 100%	Categorical	25 patients arriving for diagnostic service appointements (ongoing perspective)
What are the vital few reasons for not being financially cleared?	Reasons patient not financially cleared	Reasons patient not financially cleared (defective elements required for financial clearance)	N/A	Categorical	25 patients arriving for diagnostic service appointements (ongoing perspective)
What is the appointment demand by payer?	Appointment demand by payer	Number of appointments handled by PreCert	N/A	Categorical	Total number of appointments made Jan 1- June 30
What is the dollar value of denials?	Denials	Projected collection of 10% of charges (claims adjusted off w/o payment)	Hospital visit charges (current adjusted) x 10%	Categorical	All claims adjusted off w/o payment Jan 1-June 30

Notes:

What to Measure		Operational Definition		Data Collection	
Question to Answer	Measure Name	Words	Formulas	Data Type	Sample Size
What is the incidence of patients arriving with no orders?	Patients w/o orders	Percentage of patients presenting to C1200 with no order readily available	(# patients w/o order) / (total # registrations) x 100%	Categorical	Total number of registrations made Jan 1-ongoing

Notes:

Service

What to Measure		Operational Definition		Data Collection	
Question to Answer	Measure Name	Words	Formulas	Data Type	Sample Size
How many vendor contracts were executed by department?	Number of contracts	All vendor contracts executed	Sum of contracts	Categorical	All
What % of vendor contracts are in repository by department?	Number of contracts in repository	Vendor contracts with a copy in central repository	Sum of contracts in repository/sum of contracts	Categorical	All
How many BA relationships are there in vendor agreements?	Number of BAs	Vendor contracts that fit the definition of Business associate per HIPAA regulations	Sum of BAs	Categorical	All
What % oif vendor BAAs were appropriately executed?	Percentage of BAAs	Vendor contracts that fit the definition of Business Associate with an appropriately executed BAA	Sum of executed BAAs/Sum of BAs	Categorical	All

What to Measure		Operational Definition		Data Collection	
Question to Answer	Measure Name	Words	Formulas	Data Type	Sample Size
What % of vendor BAAs are in repository?	Precentage of BAAs in repository	Appropriately executed BAAs in central repository as a percentage of total BAAs	Sum of executed BAAs in repository / sum of BAAs	Categorical	All

Manufacturing

What to Measure		Operational Definition		Data Collection	
Question to Answer	Measure Name	Words	Formulas	Data Type	Sample Size
Is food portioned correctly?	Portioning accuracy	Percent of times food is portioned correctly vs. over portioned or under portioned	correct / total over / total under / total	Categorical	82 locations (21 no data)
What is the cost of improper portioning?	Cost of portioning	Is there a cumulative increase in food cost due to incorrect portioning	actual cost / recipe cost	Categorical	82 locations (21 no data)
Does the line of business effect portioning?	LOB	Is there a relationship between the LOB and portioning accuracy	% correct per LOB	Categorical	82 locations (21 no data)
What factors effect portioning?	Contributing factors	What is the ratio between various factors that effect correct portioning	factor / correct factor / over portion factor / under portion factor	Categorical	82 locations (21 no data)

Notes:

Steps for Collecting Data and Constructing a Data Collection Plan

By being focused and specific, questions clearly indicate what data needs to be gathered. It is also much easier to get others to help collect data if they believe that you know precisely what you are looking for and that you are going to do something with the information and results.

1. Establish Data Collection Questions: Ensure that the questions are focused and specific. Ask:
    a. What information is needed?
    b. What is the data-collection goal?
    c. What process or product will be monitored to collect the data?
2. Decide What to Measure: Decide how data will be communicated and analyzed. Ask:
    a. What data is needed, and what type?
    b. What measure will be used?
    c. What is the operational definition of each measure?
    d. What will be computed and how will it be displayed?
    e. Summary statistics
    f. Performance statistics
3. Decide How to Measure: A key element in evaluating data is the sensor. With the growth of technology, the sensors of many measurement systems have been expanded to include:
    a. Recording
    b. Data processing
    c. Comparing
    d. Actuating
4. Define Collection Points: Ideally, set up data collection at a point in the process where all the data can be collected with minimum disruption.
5. Understand Data Collection Bias: Who in the process can supply the data with minimum effort and error?
6. Design, Prepare, Test Data Collection Methods: When designing forms:
    a. Design with the collectors in mind.
    b. Design to reduce the likelihood of an error.

Notes:

c. Capture additional information for future analysis.

d. Make it self-explanatory.

7. Collect the Data: To audit the process and validate the results:

a. Randomly review completed forms.

b. Check for missing or unusual observations.

c. Look at Variation that may signal bias.

When Should a Data Collection Plan be Used?

Data collection plans should be used during the Measure and Analyze steps when an improvement team needs to collect data. This tool helps keep the collection organized and focused.

Important Points

- List the questions to be answered.
- List the tools to be used to display the data.
- Define an operational definition for the data to be collected.
- Define the sample size for the data to be collected.
- Decide where to collect the data.

Where to Collect Data:

- Existing Records
- Direct Observation
- Personal Interviews
- Telephone Interviews
- Mail Surveys
- On-line Surveys

Notes:

Pitfalls to Avoid

- Do not collect data that already exists.
- Be professional with data collection techniques and collect data properly.

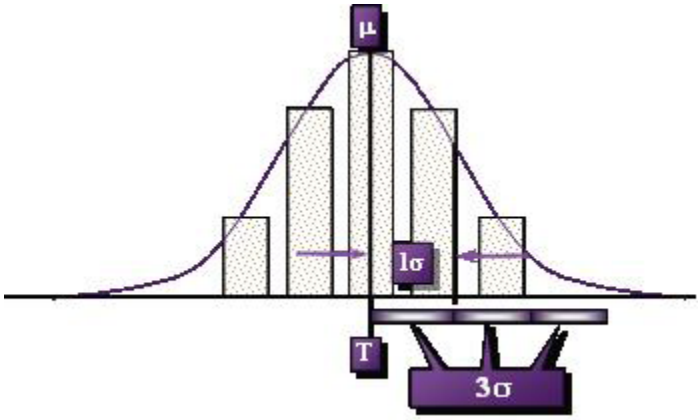
The Normal Distribution

The Normal distribution is an example of a continuous random variable distribution. There are many different Normal distributions. A Normal distribution can be identified by specifying two numbers: the mean and the variance or standard deviation.

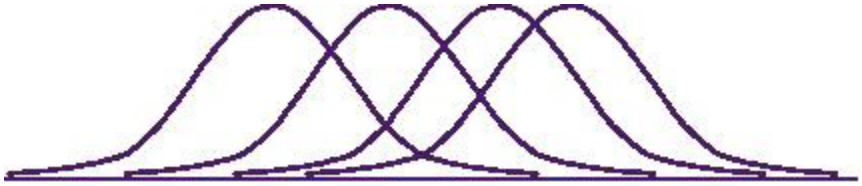
If X is a Normal random variable with mean μ and variance σ2, then the density

function Y is given by a Normal distribution which is sometimes called a bell curve. It is symmetric.

The total area underneath a distribution curve equals a probability of one. The area between two different values is the probability that the observed value will fall between those two values.

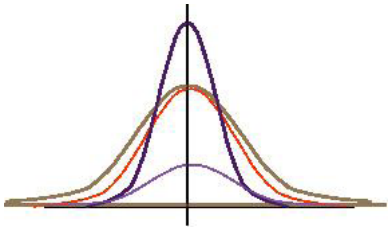


The Relationship Between The Normal Distribution, Mean, and Variance



Four normal distributions that have the same variance but different means.

Notes:



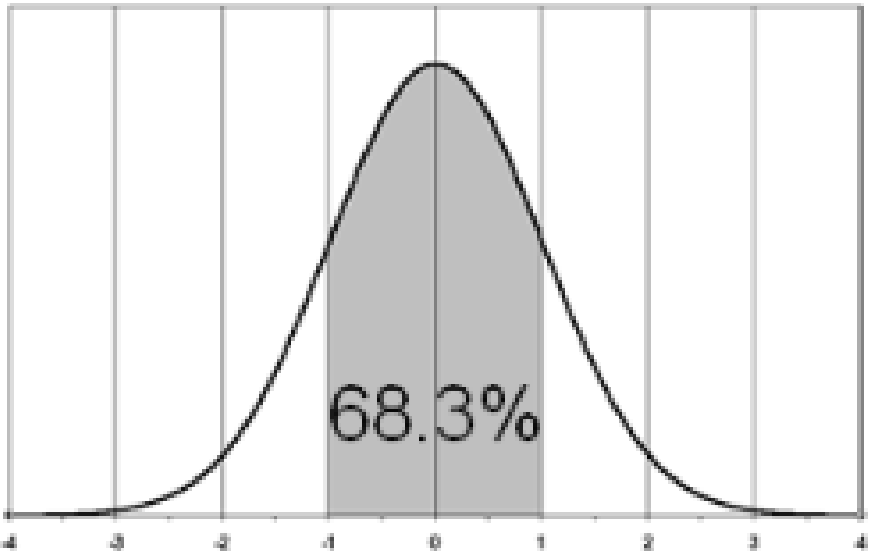
Four normal distributions that have the same mean but different variances.

Standard Deviation

The standard deviation is a measure of performance and is derived from the data.

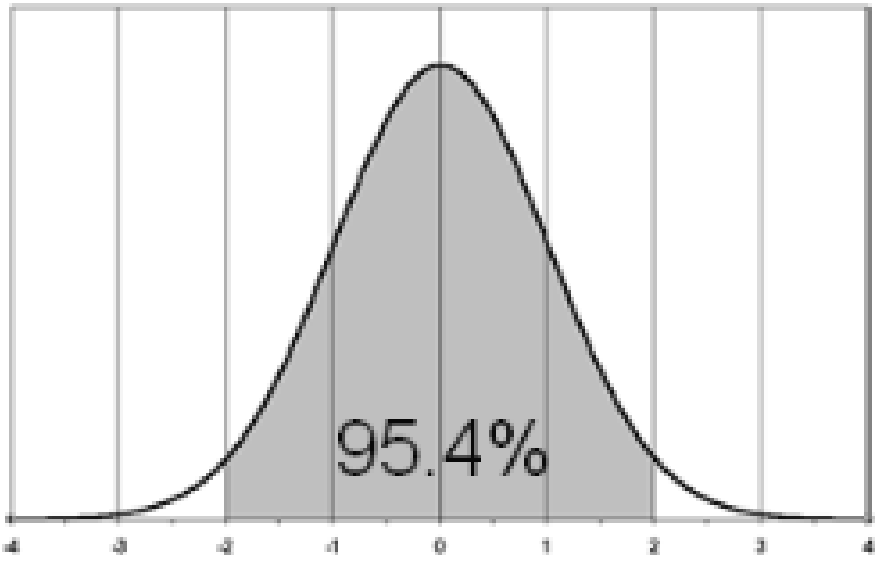
The larger the standard deviation, the fewer units of standard deviation can fit between the mean and a specification limit.

68.3% of data fall within one standard deviation of the mean

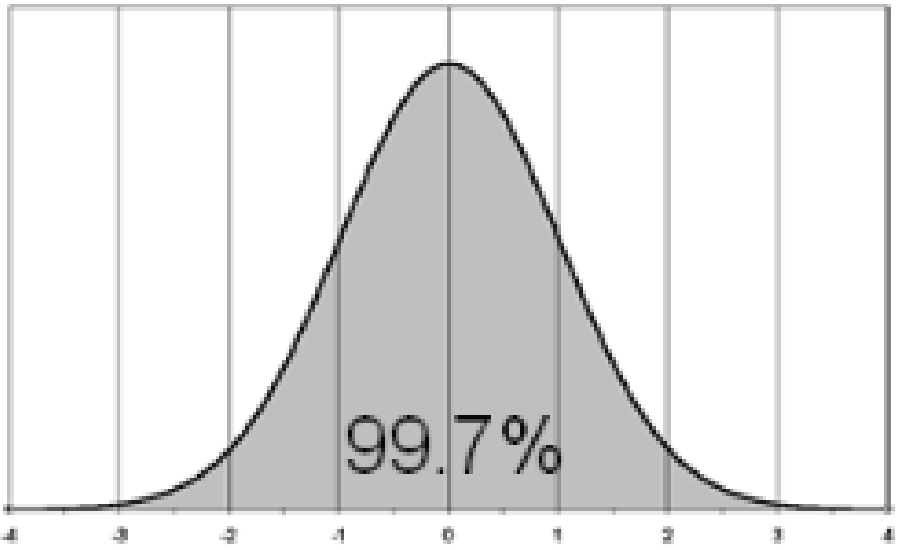


Notes:

95.4% of data fall within two standard deviations of the mean.



99.7% of data fall within three standard deviations of the mean.



Sigma Level

Sigma level describes the distance of the mean of a theoretical curve (Normal) to a specification limit. The distance is measured in units of the standard deviation(s). It is also known as the Z-value.

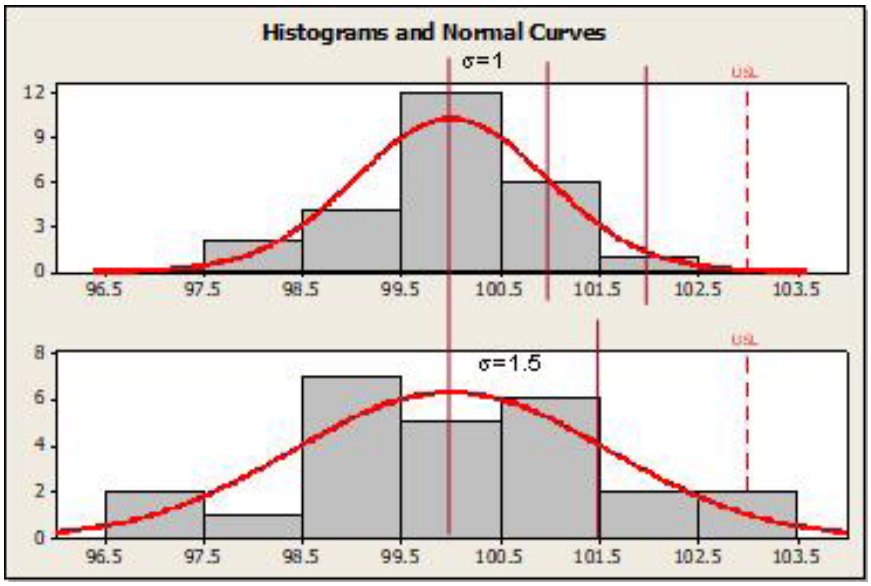
When considering standard deviation and sigma level, for a given mean and specification limit, the greater the standard deviation, the lower the sigma level will be.

Too much variation creates poorly performing processes.

Low Sigma Level = High Variation

Notes:

Sigma is a Way to Measure Performance



Both distributions have a mean of 100, and Upper Specification Limit (USL) is 103. The standard deviation for the upper data is 1, and for the lower data 1.5. The sigma level for the upper graph is 3, and for the lower graph 2.

The Ideal Unit of Measure

- Understandable: This is seldom a problem at the technological level, where the meanings of the words have been highly standardized. However, many units of measure at the managerial level involve terms which lack standardized meanings, e.g., world-class quality. Local dialects may be understood by insiders but not by outsiders, e.g., on-time arrival. Any such vagueness or confusion becomes a natural source of divisiveness. Those who lack understanding of the unit of measure become suspicious of those who possess that understanding.
- Basis for Decision Making: One purpose of measurement is to provide factual assistance for decision making by a diverse number of people. The greater the validity of the measurement concept, the greater the likelihood of securing a meeting of those minds.
- Applies Broadly: Measures of quality features are widely used as a basis for comparative analysis. You will need answers to questions such as: Is your quality getting better or worse? Are you competitive with others? Which one of your operations provides the best quality? How can you bring all operations up to the level of the best? Units of measure that have broad applicability can help you answer such questions.
- Conducive to Interpretation: Identical numbers can nevertheless result in widely different interpretations. What is critical is whether the units of measure have been defined with adequate precision. Is the measure operationally defined?
- Economical to Apply: A balance must be struck between the cost of

Notes:

making evaluations and the value of having them. The most basic question is whether it is worthwhile to measure at all. If so, then the next question relates to precision of measurement. The precision needed is whatever enables you to make valid decisions from the data. To go beyond this adds cost without adding value.

- Compatible: Measurement of quality is wonderfully simple if a ready-made instrument exists that you can plug in to read the result in terms of the unit of measure. Such simplicity is widely prevalent at the technological level of the pyramid of units of measure. However, as quality grows in importance, you are faced with creating many new units of measure.
- Measurable: Some quality features seem to stand apart from the world of physical things. Quality of service often includes courtesy as a significant quality feature. Even in the case of physical goods, there are quality features, such as beauty, taste, aroma, feel, or sound. The challenge is to establish units of measure.

Notes: