Introduction

Module 1

6σ

Green Belt Program

Introduction

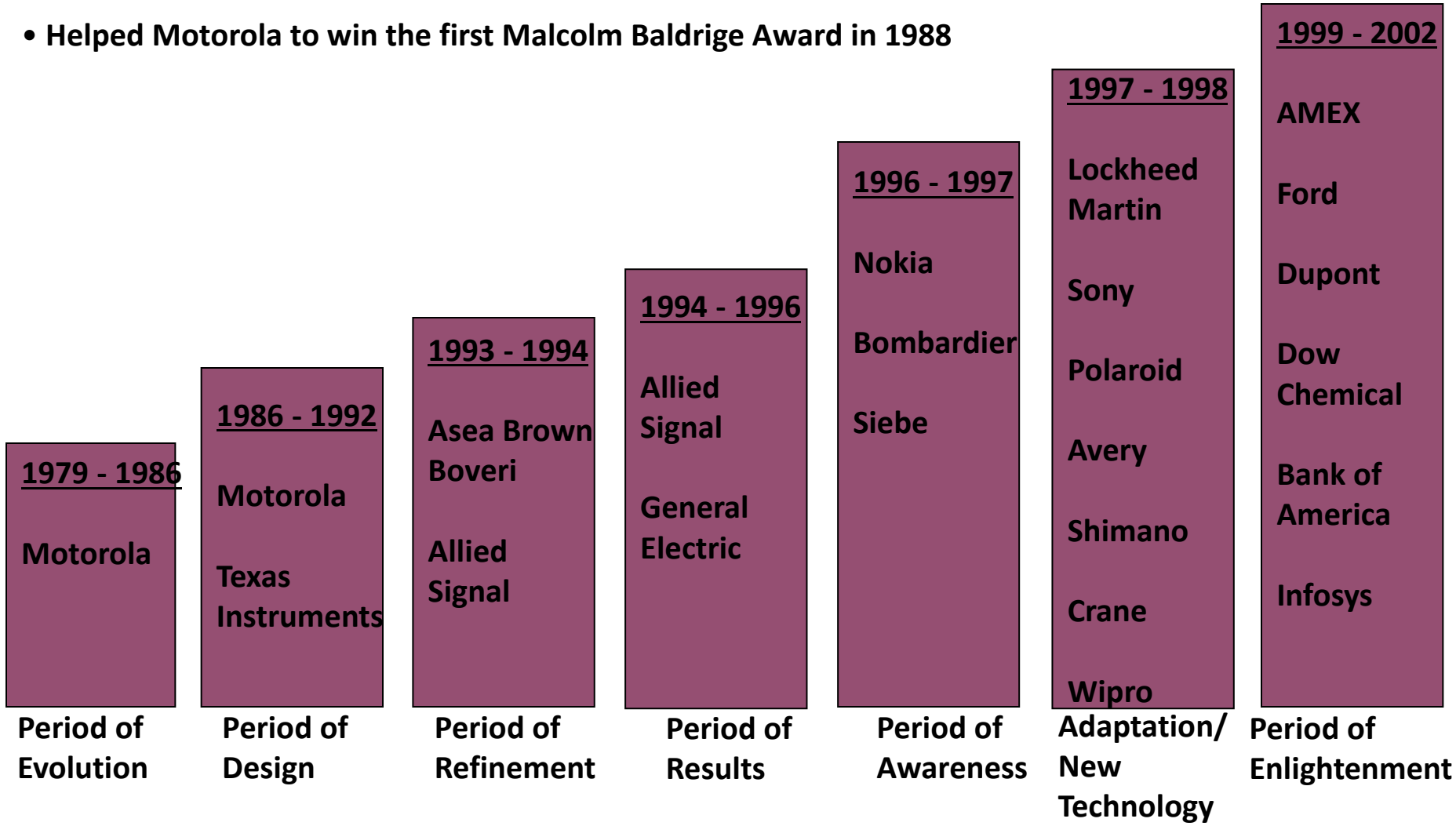
Module 1

OBJECTIVES

- Understand and Appreciate the application of Six-sigma methodology for process or product or service improvements
- Improve Analytical skills for Problem Solving
- Good Grasp on Six-sigma tools for Process Diagnosis
- Apply DMAIC to your project as the primary process improvement methodology
- Insight into Certification Process

Introduction

- Developed by Bill Smith & Mikel Harry, Motorola
- Helped Motorola to win the first Malcolm Baldrige Award in 1988



Introduction

Six Sigma: Breakthrough Improvements

Reduces Variation

Reduces Cycle Time

Reduces Defects

Leads to

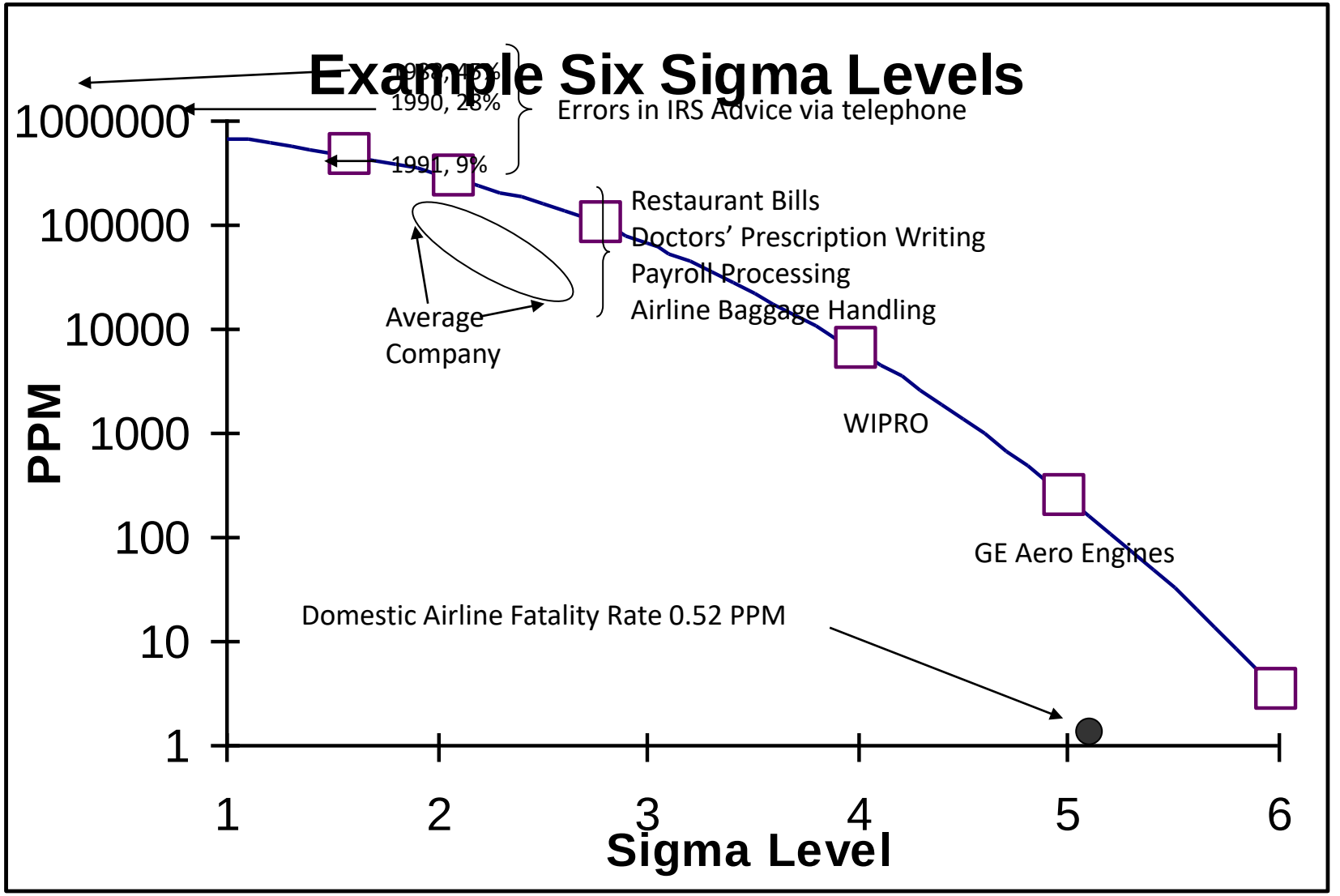
Improved Customer Satisfaction

Improved Bottom Line (Net Profits)

Introduction

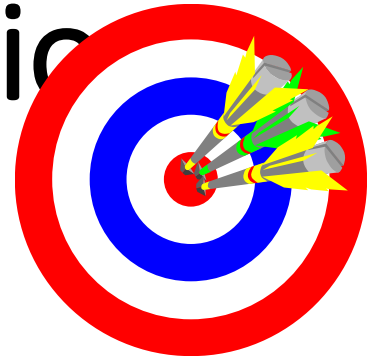
Sigma Level	PPM	Improvement
6.0	**3.4**	**68.5**
5.5	31.7	
5.0	**232.6**	**26.7**
4.5	1349.9	
4.0	**6209.7**	**10.8**
3.5	22750.1	
3.0	**66807.2**	**4.6**
2.5	158655.3	
2.0	**308537.5**	**2.2**
1.5	500000.0	
1.0	**691462.5**	**1.3**
0.5	841344.7	
0.0	**933192.8**	**1.0**

Introduction



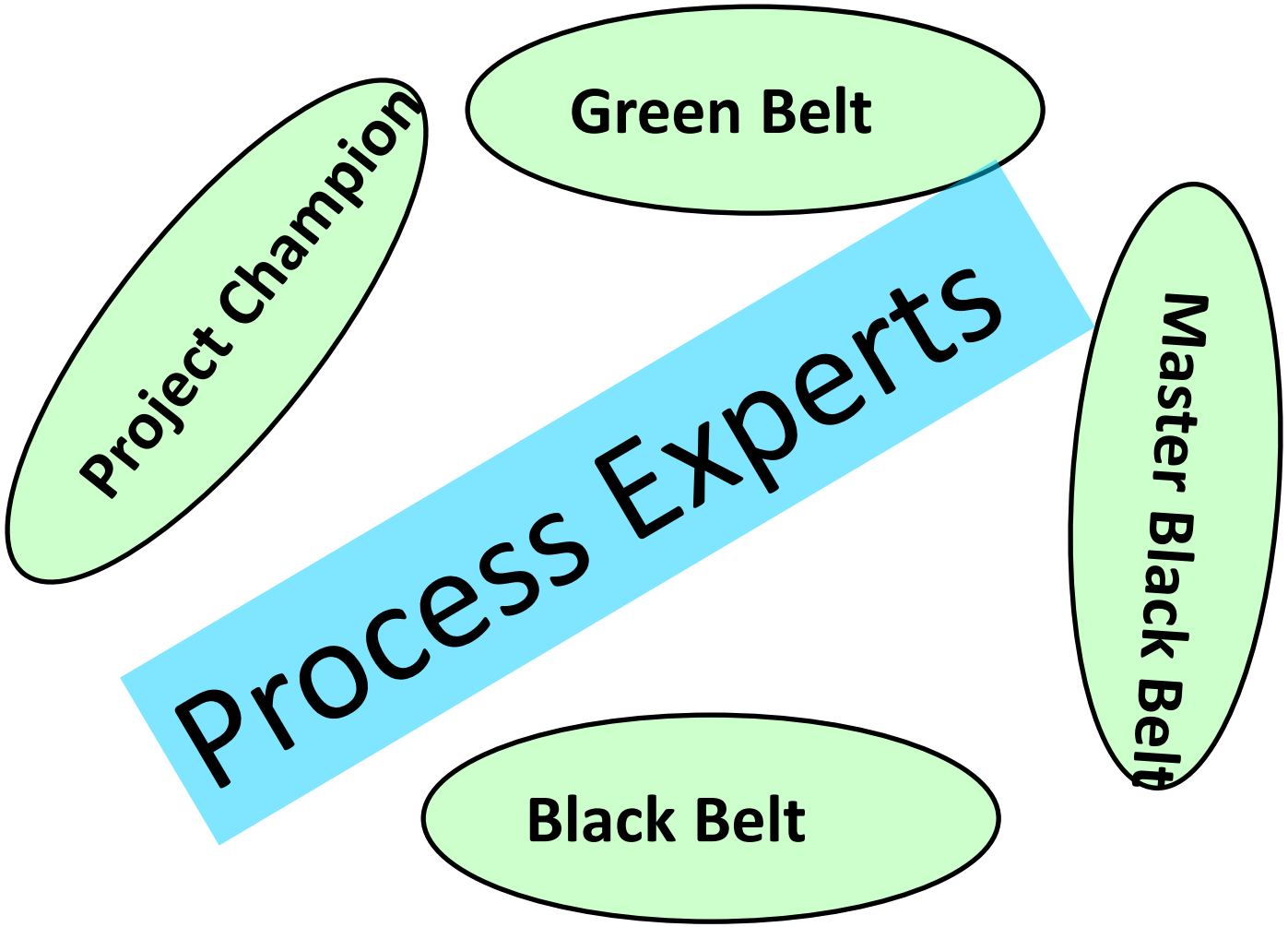
Introduction

Introductio

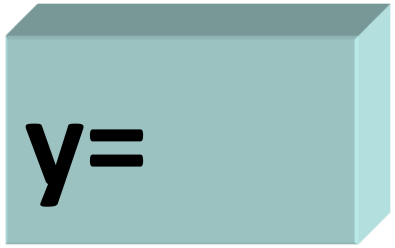


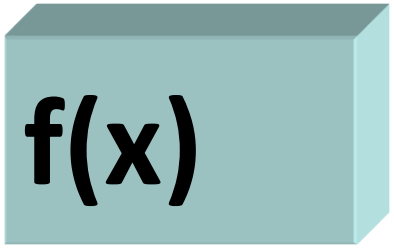
Business Problem	Statistical Problem	Statistical Solution	Business Solution

Six Sigma applies statistical tools to Business problems.
The key is data-driven decision making.


Green Belt
Project Champion
Process Experts
Master Black Belt
Black Belt

Introduction



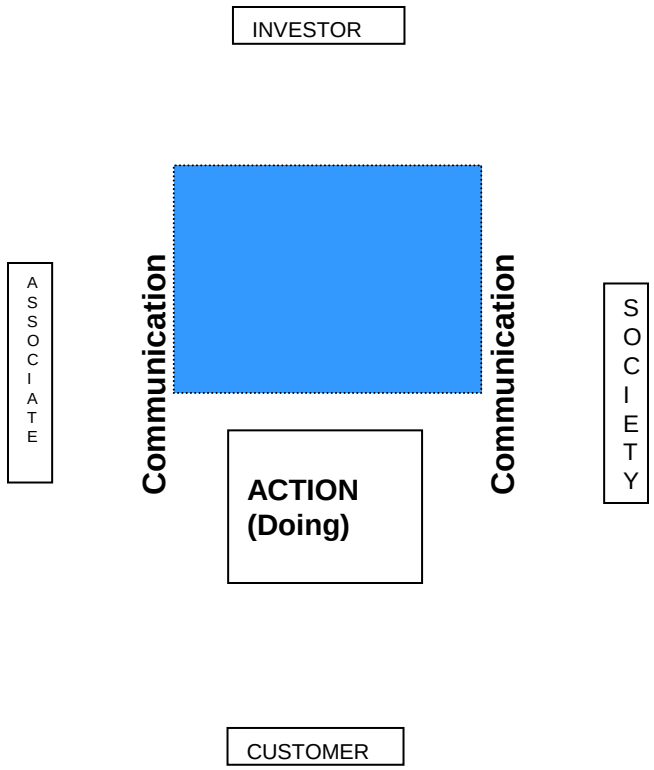


To get results, should we focus our behavior on the y or x ?

- **y**
- **Dependent**
- **Output**
- **Effect**
- **Symptom**
- **Monitor**

- **$x_1, x_2, \ldots . x_n$**
- **Independent**
- **Input-Process**
- **Cause**
- **Problem**
- **Control**

Y = f (Man , Machine , Method , Measurement, Material)

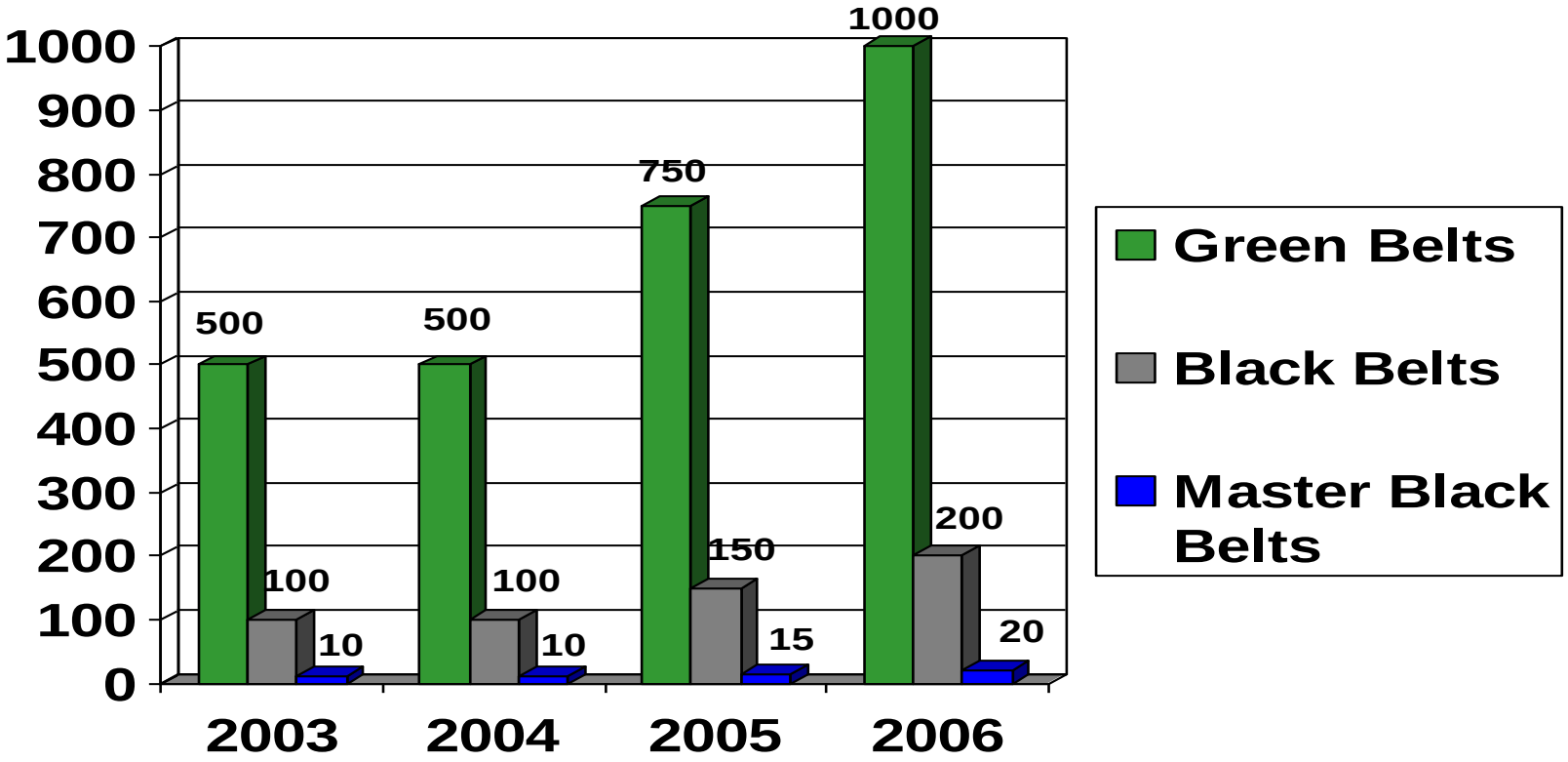


Vision – Measure & Lead

6σ

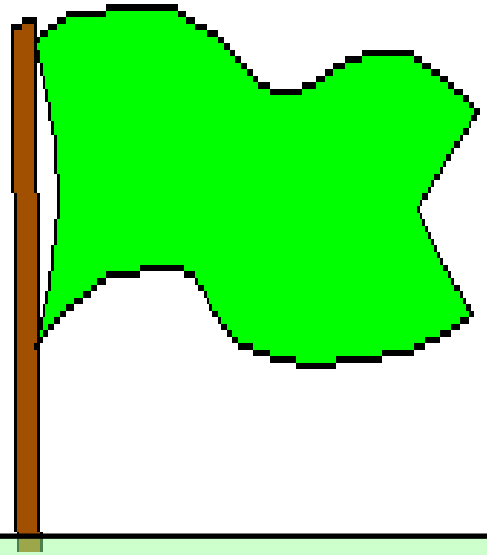
Six Sigma

Communication



- 6σ Certified Resources (6σ Associates, GB, BB, MBB)
- 6σ Process Improvement Projects resulting in significant Cost Benefits and Stakeholder Delight

Green Belt



Projects

Successful completion of 1 Six-sigma Process improvement Project

Training

- 2-day : DMAIC methodology + Tools & Techniques

A-I-C-S Delight

Six Sigma

Defect Reduction & Cycle Time Improvement

Cross Functional Processes

Personal Six-Sigma

Software Quality Improvement

DMAIC Methodology

CFPM using DMAIC Methodology

DFSS (DMADV) Methodology

Six Sigma Initiative provides methodology for Quantitative Process Management

Process Problems

Accuracy

Problems of Mean (or Average Performance)

Off-Target

Center Process

Accuracy & Precision

Problems of Mean & Dispersion

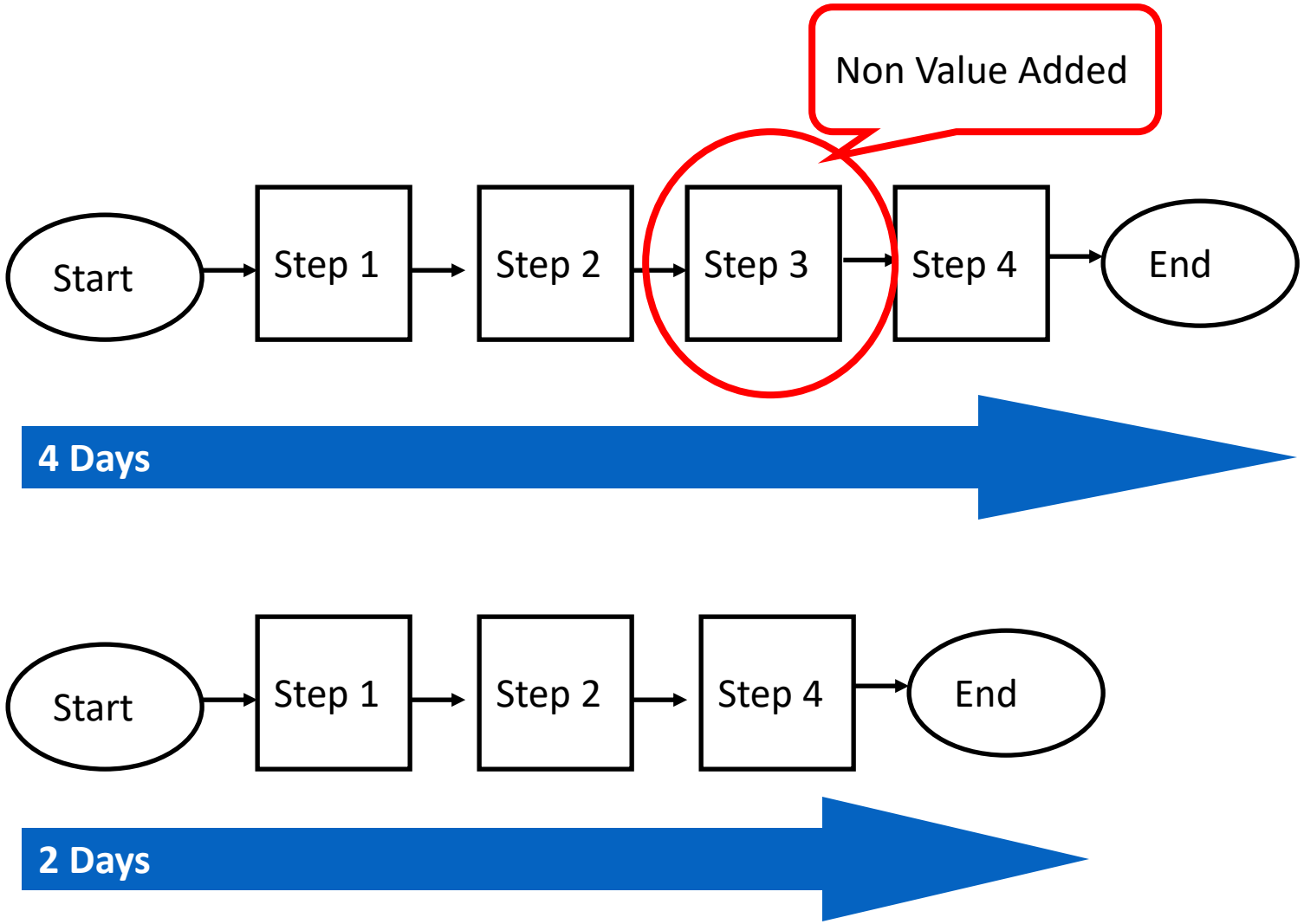
On-Target (Precise & Accurate)

Precision

Problems of Dispersion (or Performance Variation)

Variation in Performance

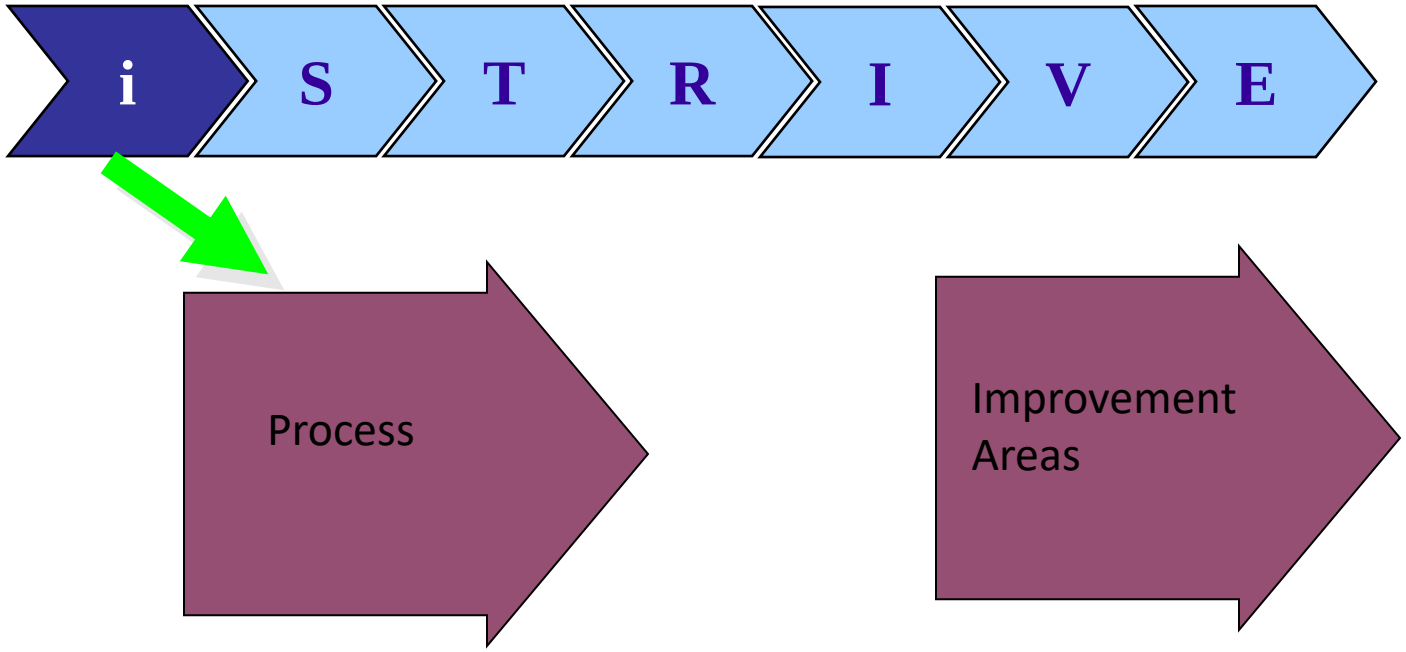
Reduce Spread


Non Value Added
Start
Step 1
Step 2
Step 3
Step 4
End
4 Days
Start
Step 1
Step 2
Step 4
End
2 Days

Identify

Module 2

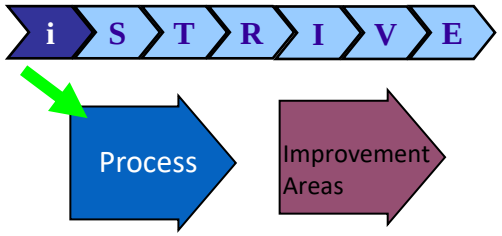
6σ



- Identify Processes within/ between Circle(s)
- Identify High Level Business Process Map
- Identify Circle Measures Y's (both Process ML's and Circle Level Measures)

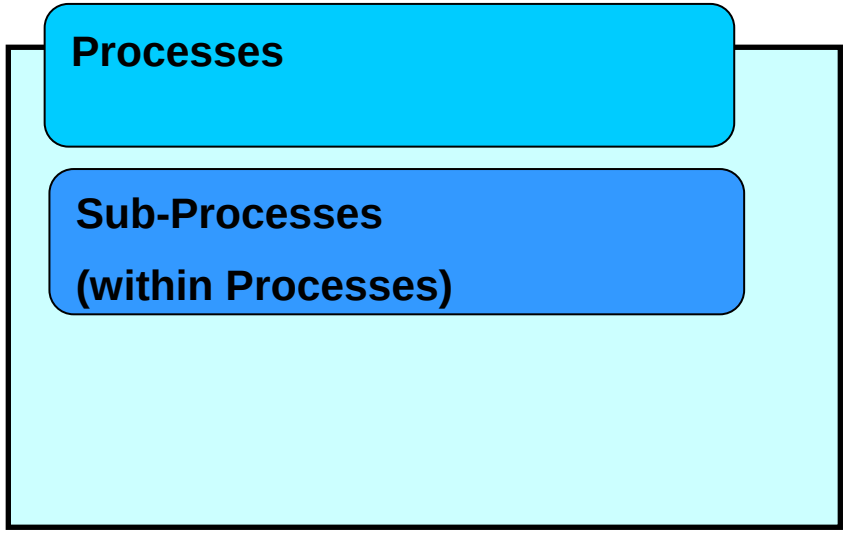
- Establish relation between Circle Level Measures and Circle ML's
- Identify Targets for Y's
- Baseline Y's (As is Performance)
- Prioritize Ys for improvement

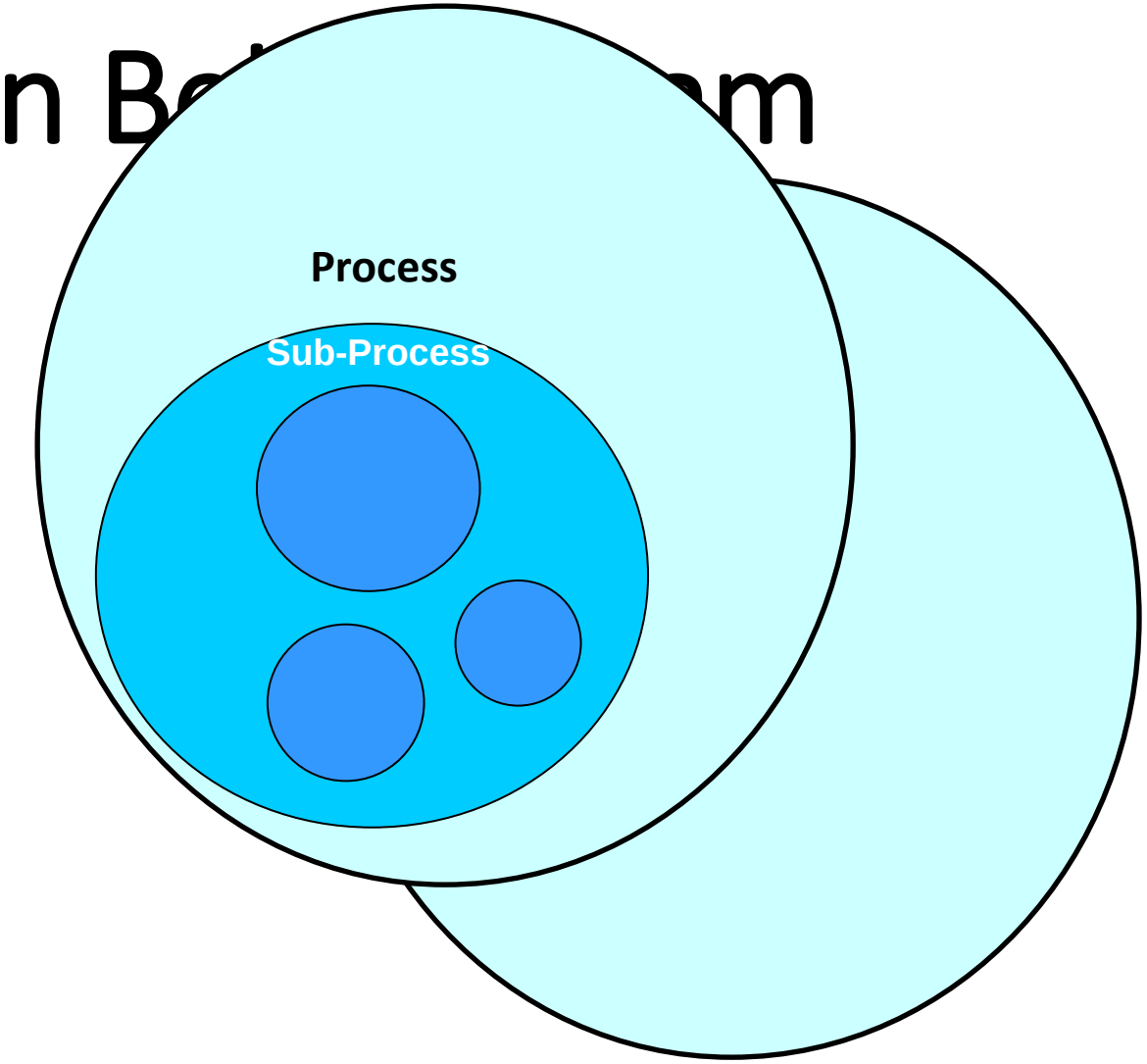
Identify the Processes to be improved

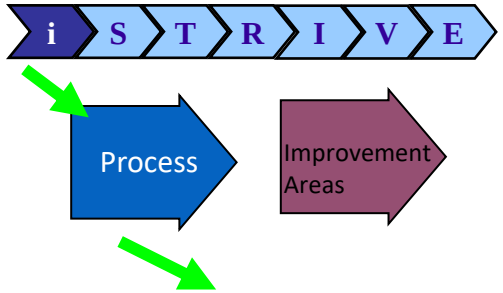


Green Be m

Processes
Within and Between Circles

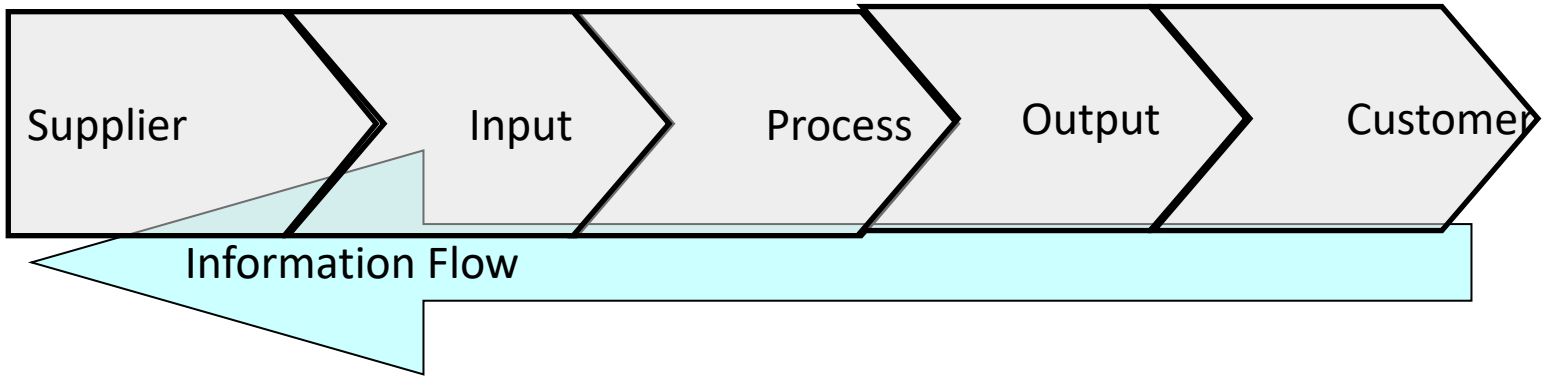


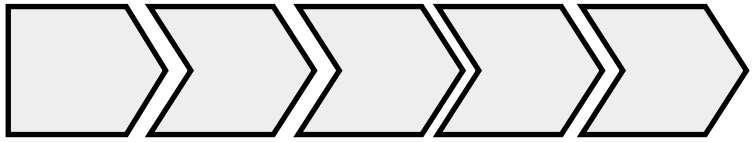




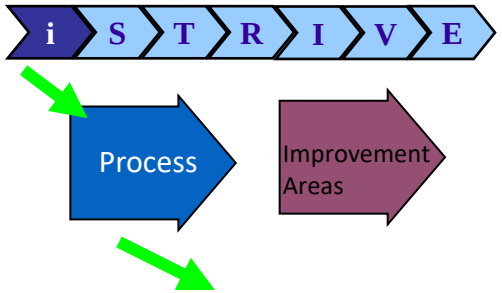
Green Belt Program

Develop a high-level Business Process map





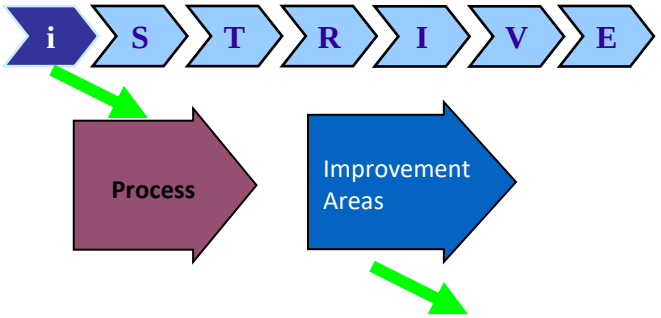
- Draw a macro picture of the process
- Identify the Process Stakeholders (A-I-C-S)
- Establish Clear Boundaries - What is required to start the process ? Who should provide that ? What would be delivered by process to Customer ?



Measures
Green Belt Program

1. **Outcome Measures – Results Framework**
    1. High-level measures on ***what*** you are doing - assess overall performance in the areas you are measuring
    2. External in nature and are most closely tied to outputs, customer requirements, and business needs for the process.
2. **Build Measures – Process Framework**
    1. Measures that ascertain ***why*** a process is not performing up to expectations.
    2. Internally focused and are usually associated with internal process steps and inputs received from suppliers.

Measure the right things



Green Belt Program

Circle Performance – CML (∑Y)

Business Performance

∑Ys

MLs of SSUs

MLs of SBUs

Ys

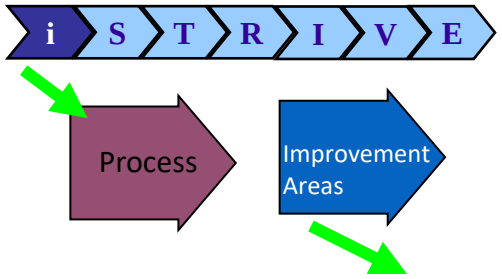
ML of Processes

ML of Service Offerings (Marketing & Delivery)

Process Measures

$\sum Y = f(Y_1, Y_2, \ldots Y_i \ldots Y_n)$

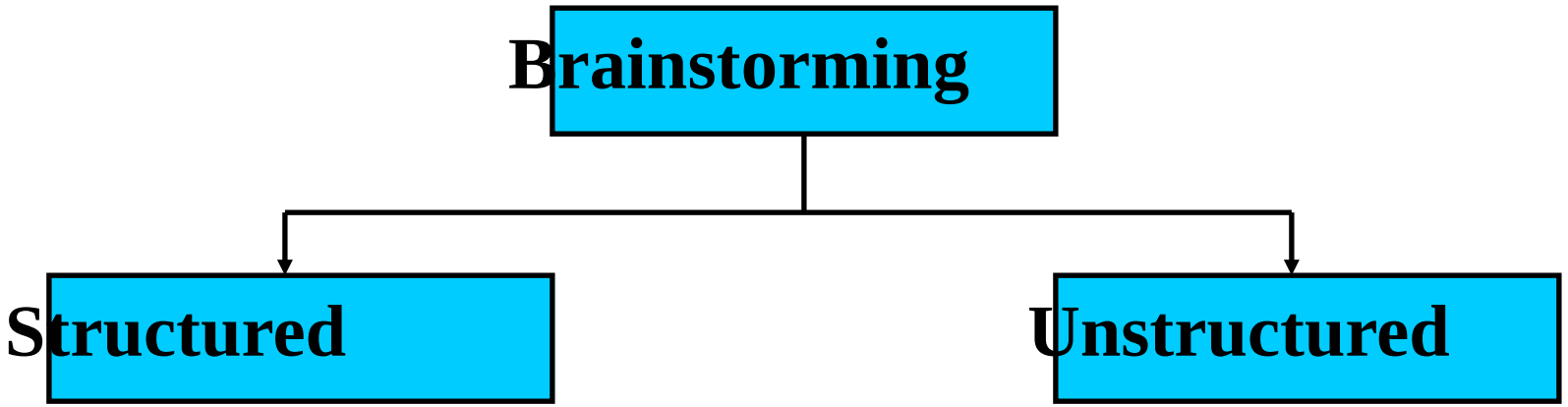
Green Belt Program



Green Belt Program

Identify the Improvement Opportunities

- Evaluate the current process capabilities
- Identify the Gaps in Performance
- Prioritize the Improvement Areas



- A creative thinking activity that involves suspending judgment and analysis
- Solicit maximum number of ideas and suggestions from a team without regard to the quality of the ideas or suggestions

Guidelines for Creativity

- Think : Use your mind continuously and improve your thinking power
- Record : Develop the ability to write and organize your thoughts.
- Devise : Develop and implement the improvements yourself.
- Calculate : Look ahead and become aware of raw cost.
- Action : Be flexible in your action.

Questions to be answered

1. What are the Processes & their Deliverables?

2. Who are the Stakeholders?

3. What are the measures specific to the process/service offering, and circles ?

4. How is the Performance Index (PI) – Ys of each Process, and the Circle PI (∑Y)

5. What are the Gaps & Opportunities?

6. What are the key processes to be improved (critical to performance –∑Y) ?

Deliverables - Phase level

1. High Level Process Map of Circle

2. Baselining of Circle PI (∑Y) and Process PI (Y)

3. Key Circle Measures Ys influencing ∑Y

Tools:

Do's

1. Identify the Macro-level process map with all its sub-processes

2. Collect data on all Measures

3. Review the relative importance of Circle Level measures over the Process PI

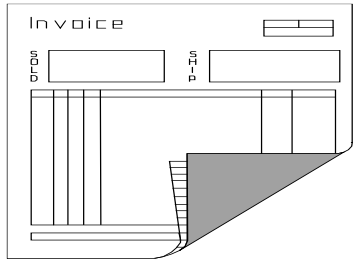
4. Baseline Measures

5. Identify the processes (Y's) to be improved

Don'ts

1. Decide on the Improvement areas based on the "Gaps" perspective alone wrto their Targets. Consider their relative influence on the overall circle Performance (ΣY)

2. Proceed without a buy-in of all key Stakeholders.



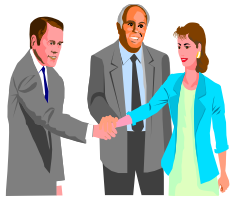
Process

Stakeholders

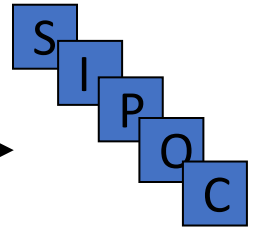
Customer

Investor

Associate

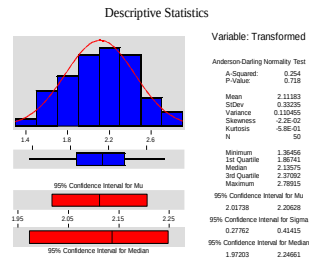


Society



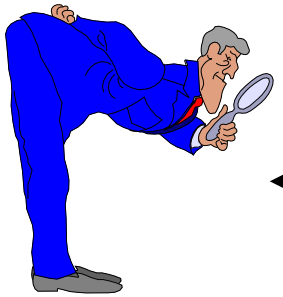
Macro level Process Map

Vision (Measures & Targets)



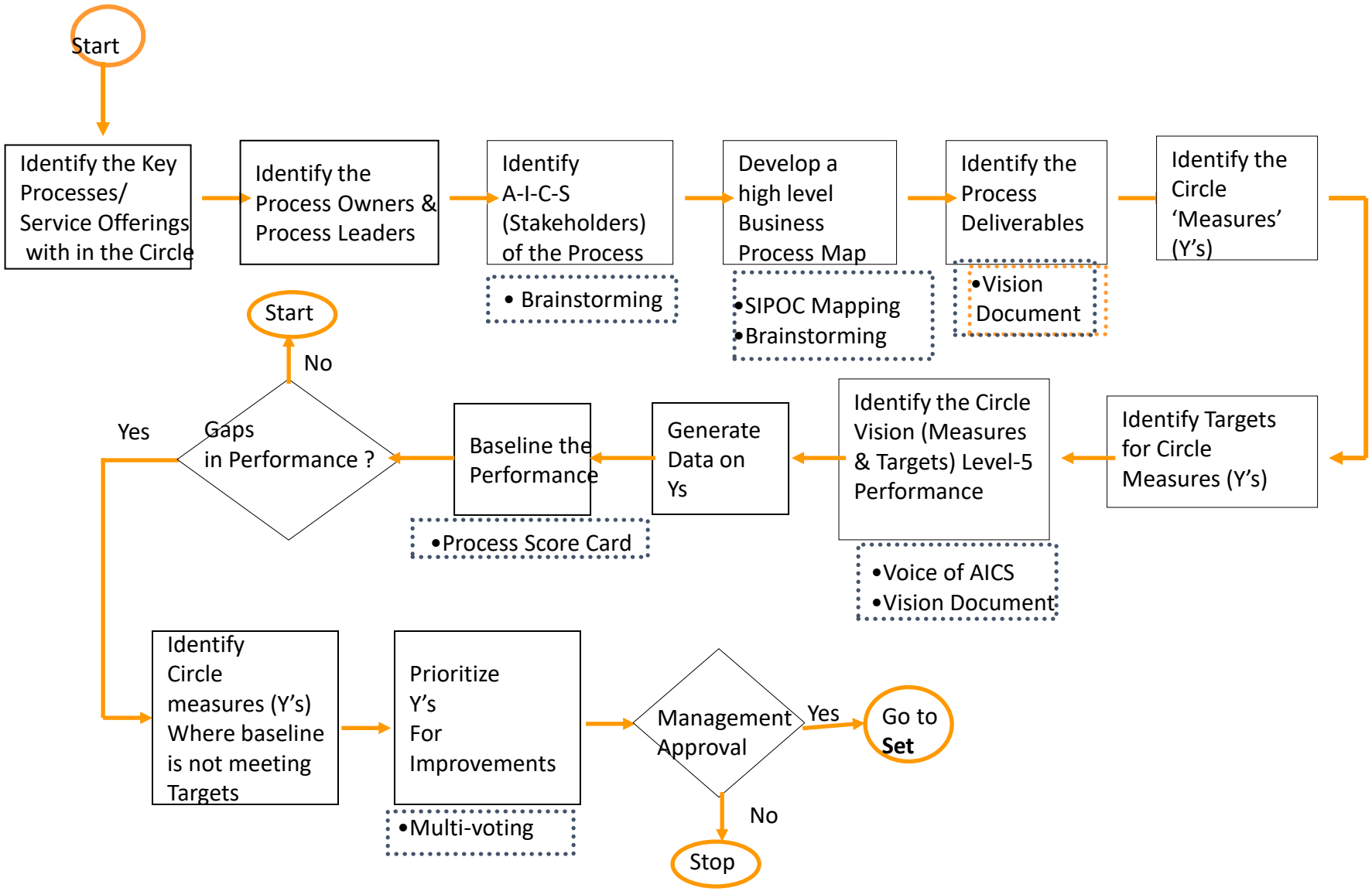
Baselining - Quantify the existing Process Performance

Gaps



Improvement Areas

Identify the processes to be improved

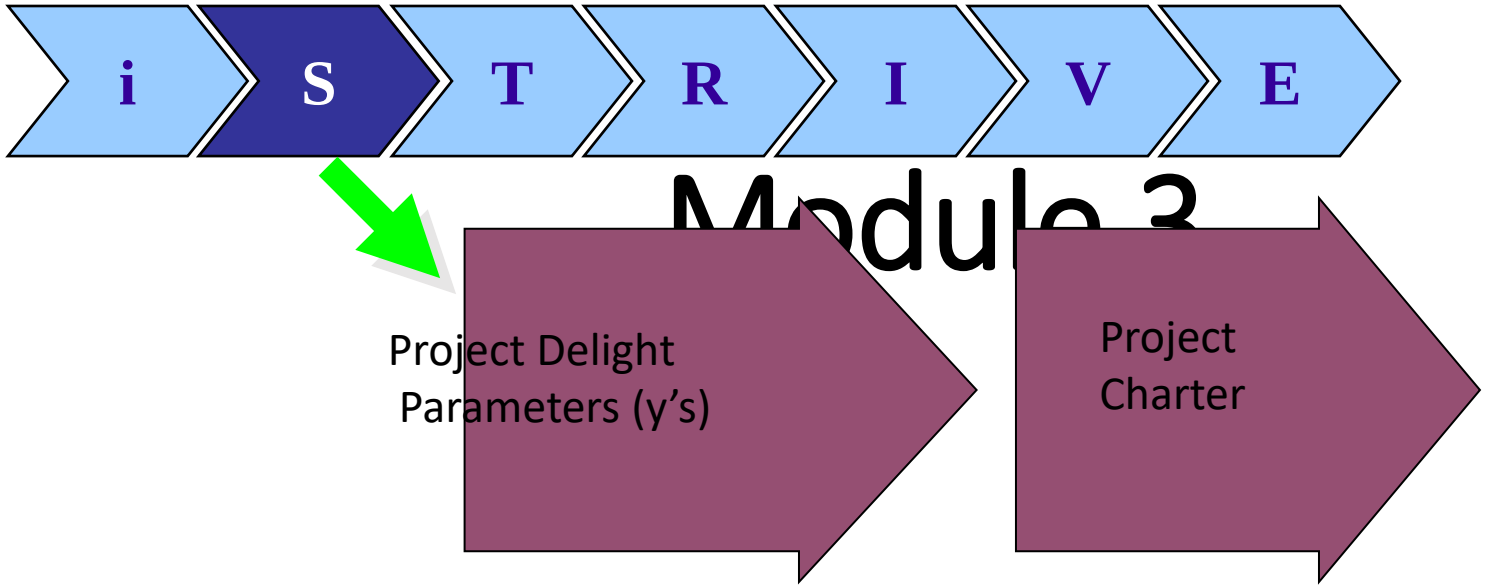

Start
Identify the Key Processes/ Service Offerings with in the Circle
Identify the Process Owners & Process Leaders
Identify A-I-C-S (Stakeholders) of the Process
• Brainstorming
Develop a high level Business Process Map
•SIPOC Mapping
•Brainstorming
Identify the Process Deliverables
•Vision Document
Identify the Circle 'Measures' (Y's)
Identify Targets for Circle Measures (Y's)
Identify the Circle Vision (Measures & Targets) Level-5 Performance
•Voice of AICS
•Vision Document
Generate Data on Ys
Baseline the Performance
•Process Score Card
Gaps in Performance ?
No
Start
Yes
Identify Circle measures (Y's) Where baseline is not meeting Targets
Prioritize Y's For Improvements
•Multi-voting
Management Approval
Yes
Go to Set
No
Stop

Microsoft Word
Document

See appendix for the case study:
“The Case of Stranded Supreme”

SET

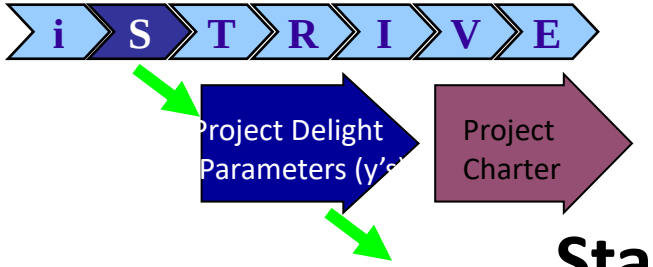
Module 3



- Capture Voice of Stakeholders (VOS) on selected Y's for improvement
- Translate VOS into y's (Process Measures, which are Candidates for Projects)
- Prioritize y's
- Select key y's for projects

- State the "Business Problem"
- Select the Project Team and define Resource requirements
- Initiate 'Project Charter' for each 'y'
- Obtain Management Approval

Set the Project Delight Parameters (y's) & Goals



SET

Stakeholder Focus

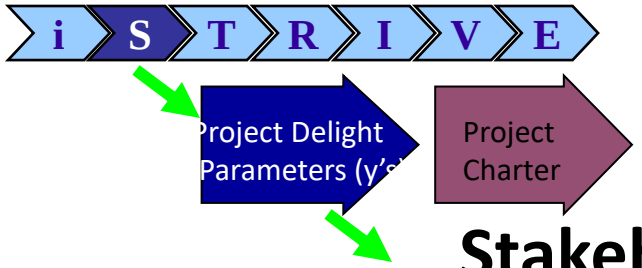
Six Sigma begins with the A – I – C – S

It is easier to define what they do not want

Process Measures - y's are defined by Stakeholder

- **Surveys**
- **Interviews**
- **Focus Group**
- **Suggestions**
- **Observation**
- **Competitive Comparisons**

Voice of A-I-C-S → y's



SET

Stakeholder Focus

AICS requirements must be understood clearly

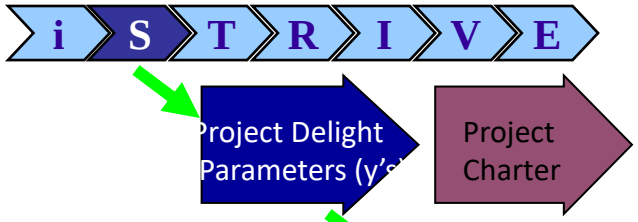
VOS is a technique to organize, analyze & profile the AICS requirements

Voice of Stakeholder Table							
S#	Who ? (A-I-C-S)	What they said ? (Voice of A-I-C-S)	What they meant ?				
			What is the Need ?	When is the Need felt ?	Where is the need felt ?	Why is the need felt ?	How is the process now ?
	Associate						
	Investor						
	Customer						
	Society						

Example – Travel Process

SET

	Associate						
1	Travel Desk (CS)	Associates do not plan in advance most of the times	Travel desk activities require a specific lead time	When last minute requests flow-in	When tickets are not available	To meet travel request needs	Not structured
	Investor						
2	Organisation/ Corporate	Management has clearly instructed to cut “excessive” Travel Costs	Reduce Travel Costs	When Travel Costs offshoot Budgeted figures	Domestic Travel	Increase in Air Travel Costs	Air Travel without prior budgets approval
	Customer						
3	Associates travelling	We keep waiting for travel confirmation & cannot commit to clients	Faster Response from Travel Desk	When an approved TR is submitted	International Travel	Time to commit to clients is important & affects Business	Total Cycle Time between Travel need identification to receipt of tickets, takes 15 days
	Society						
	nil						



SET

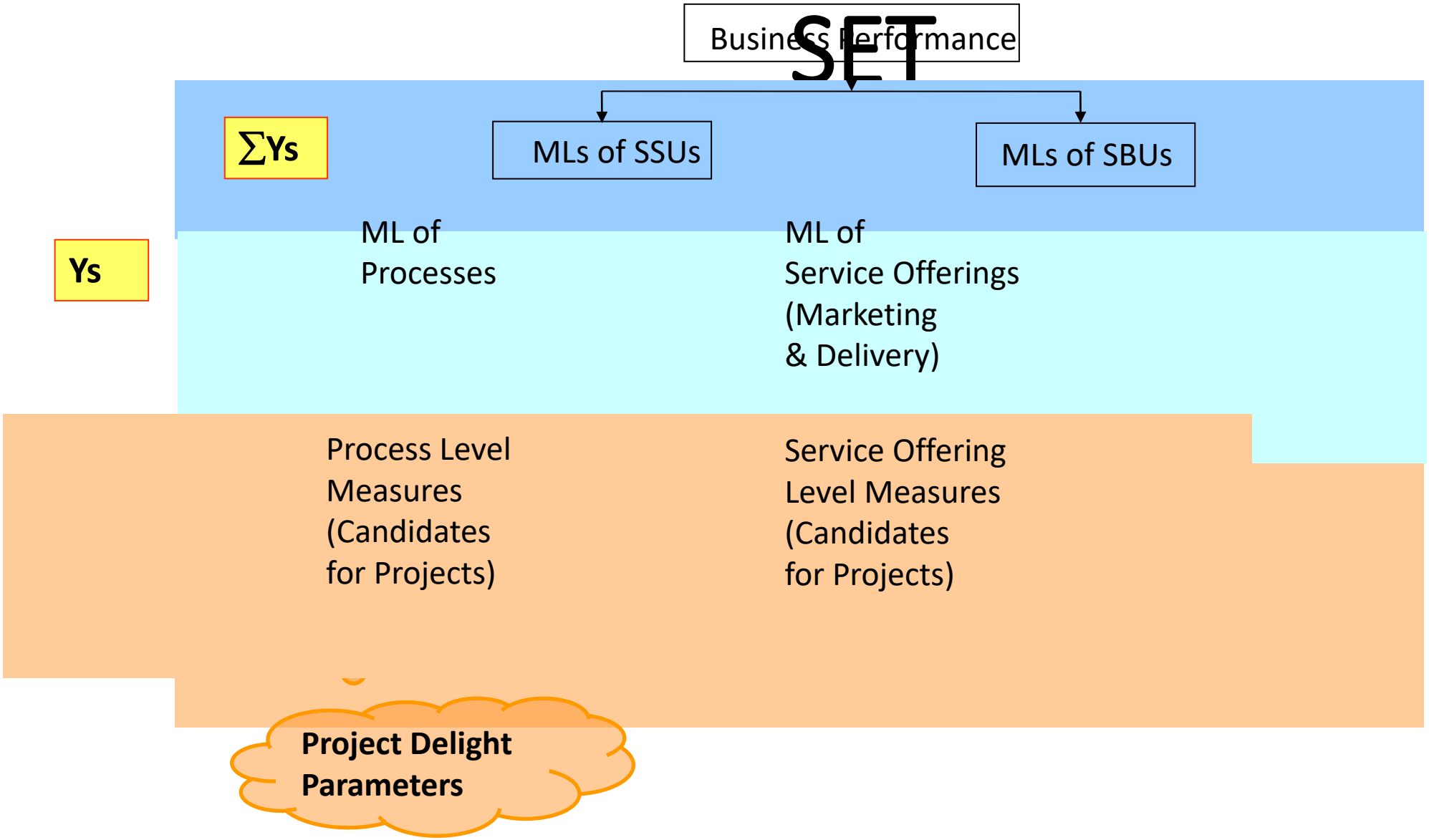
Voice of A-I-C-S	Service / Quality Attribute	Needs Statement	DP - Project y (output Characteristic)
We keep waiting for Travel confirmation And can not commit to Clients	Cycle Time : Travel Booking Confirmation / Delivery	**Status Confirmation / Tickets to be delivered to the associate as per his/her Needs (Correct) ;** Customer wants faster response (incorrect)	Timely Response

AICS needs are often high-level, vague and non-specific. These must be converted into meaningful internal goals that are assignable to functions Ex : I need Quick Response

Quick Response **Project y**

Timely Response (Project y Output Characteristic)
Time from Request to Delivery (Project y Metric)
8 hours or Less (Target)
Not greater than 2 days (Specification / Tolerance Limit

Green Belt Program



Steps in Project Selection

SET

Some Key Questions:

1.How important is it to your customer?

2.Is there a champion who feels that the project is important?

3.Is it a part of GB’s current job responsibilities & objectives?

4.Is the y (Project Delight Parameter) measurable?

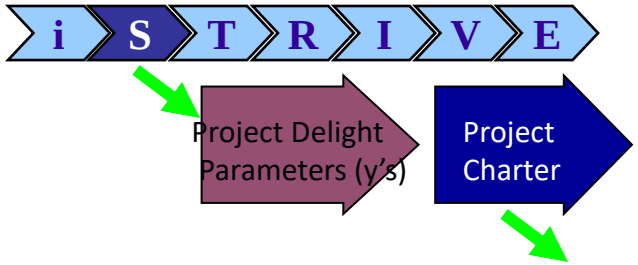
5.Is data available or easily generated?

6.Are the benefits easily measurable?

7.Is the process stabilized or at least controllable?

8.Is the scope narrow enough to finish in 4 to 6 months?

9.Does the project have a high potential for replication?

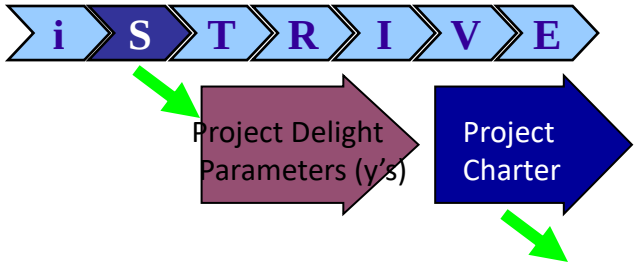


SET

Project Charter details the list of project activities and the organizational resources committed for the project.

Business Case	Problem & Goal Statement	Scope	Project Milestones	**Resources; Team Roles & Respblts.**

Project Plan documented and reviewed with Champion / sponsor

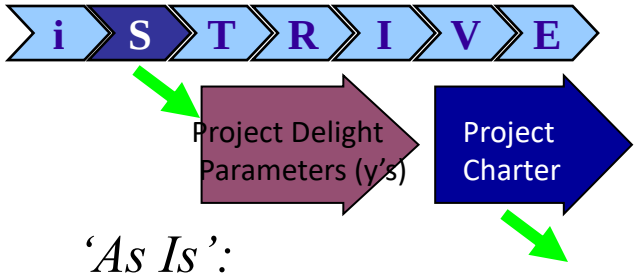


SET

Business Case	Problem & Goal Statement	Scope	Project Milestones	**Resources; Team Roles & Respblts.**

Business Case

- Background, Impact of the Problem, Need for improvement
- Business Problem linkages with the Project Delight Parameters (PDPs)
- Financial benefits / DPMO reduction / Cycle time reduction



Green Belt Program

Business Problem Statement

'As Is':

- Describes the problem as it is today
- Should not contain causes
- Should not imply solutions
- Should be as specific as possible and include measurement

'Desired State':

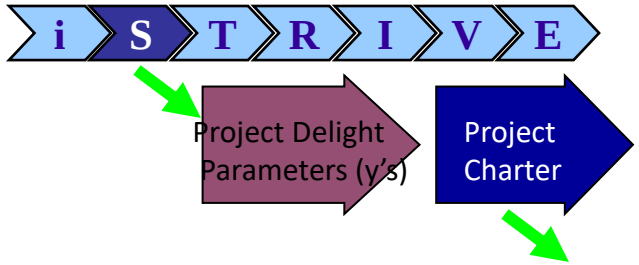
- What you want to achieve by solving the problem
- Be as objective as possible
- Be as specific as possible including measurement goal

Example:

As Is (Problem Statement): The response time for 15% of our service calls is more than 2 hours.

Desired State (Project Objective): The response time for all service calls must be 2 hours or less.

Green Belt Program

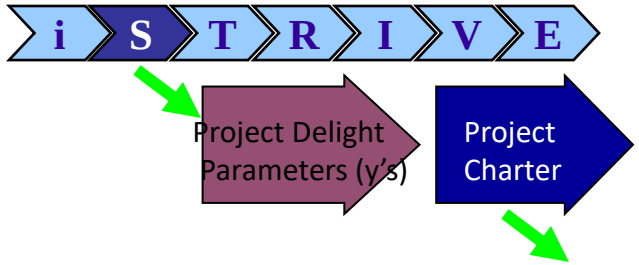


Business Case	Problem & Goal Statement	Scope	Project Milestones	**Resources; Team Roles & Respblts.**

Problem & Goal Statement

- Description of the Problem/Opportunity
- Write down the Goal / Target that you want to achieve

Green Belt Program

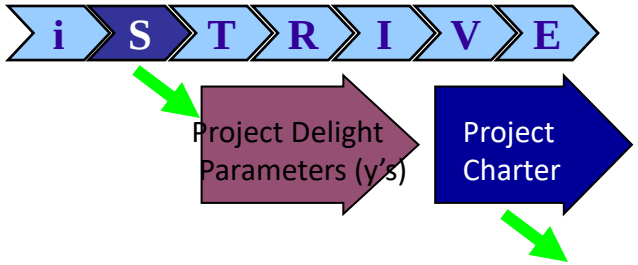


Business Case	Problem & Goal Statement	Scope	Project Milestones	**Resources; Team Roles & Respblts.**

Scope

- Boundaries of the process
- Constraints & Project Dimensions
- What is included and what is not

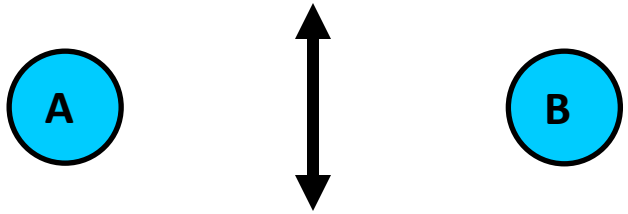
Green Belt Program



Types of Scoping?

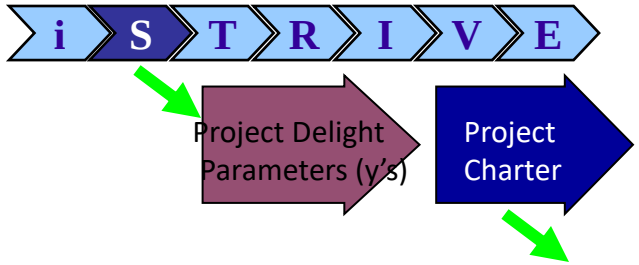
Longitudinal

(Length)

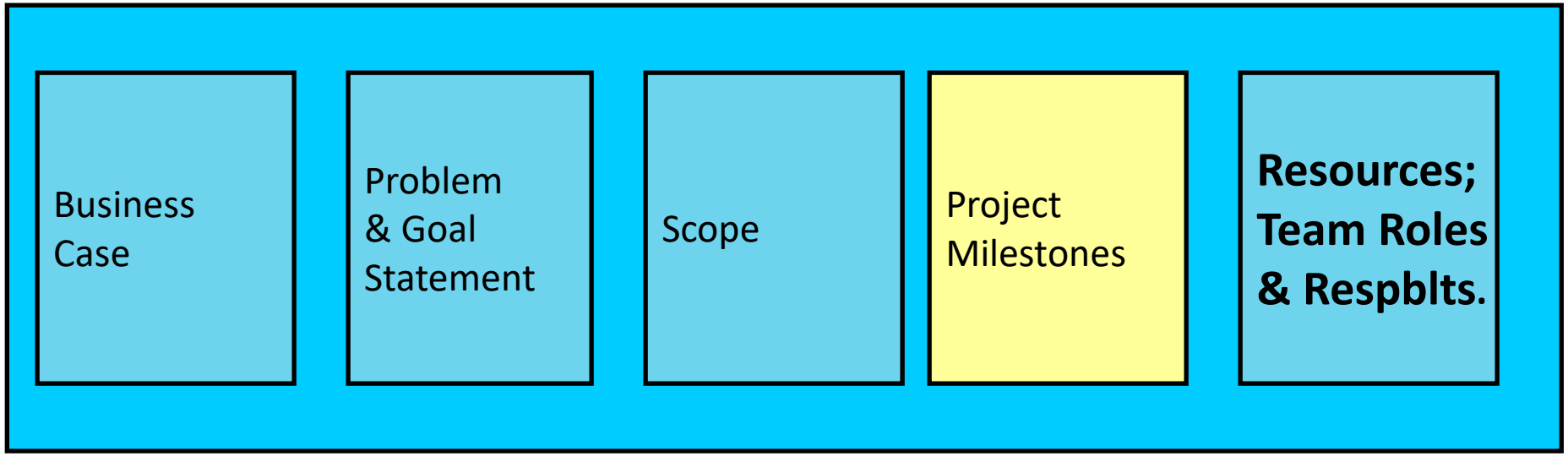


(Breadth)

Process



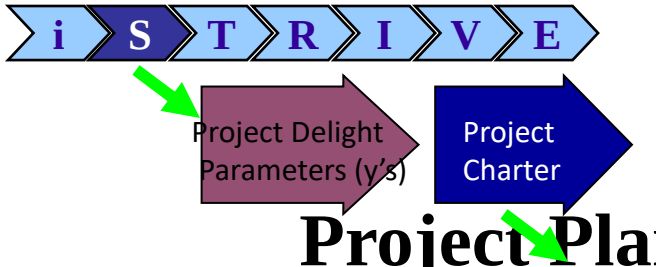
SET



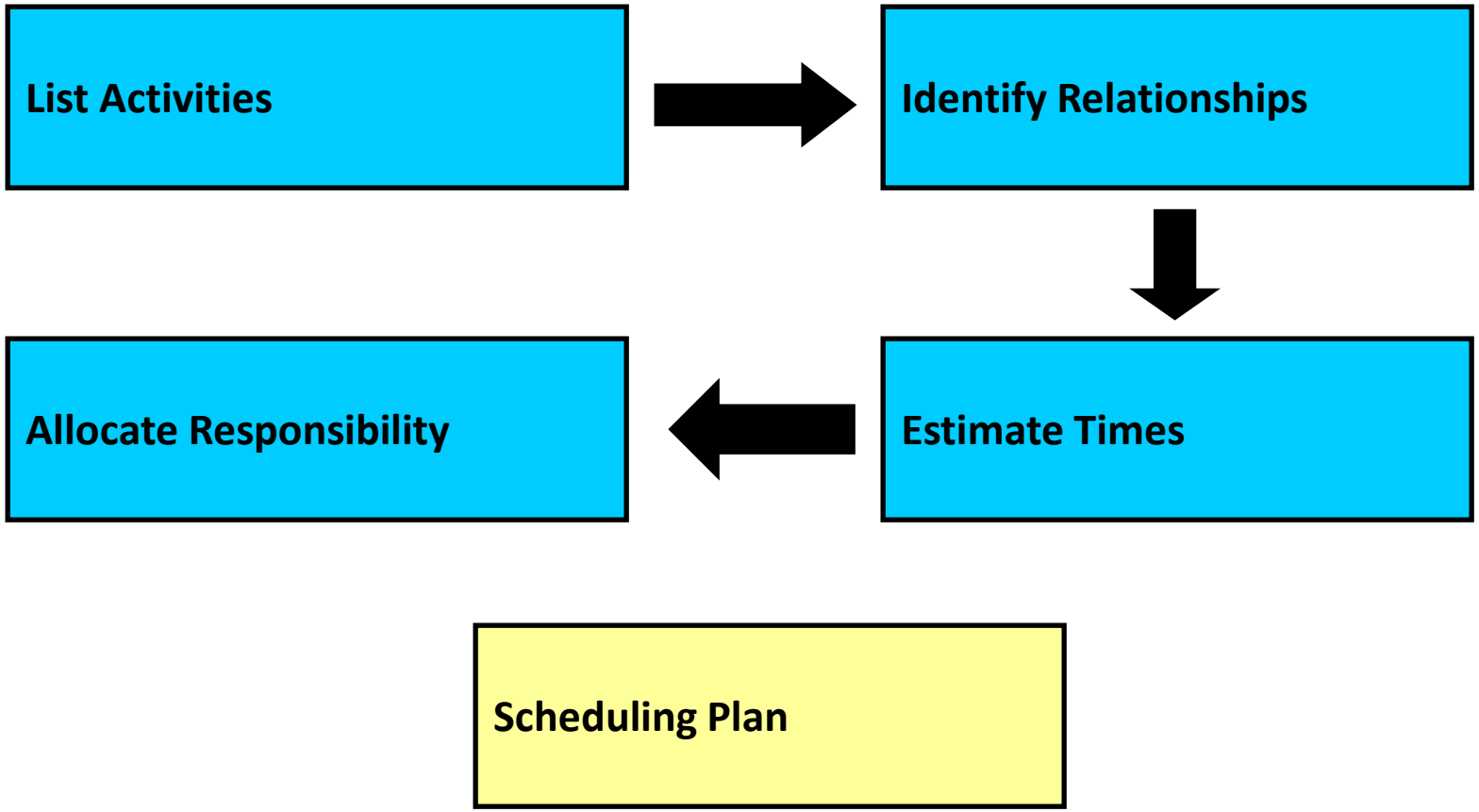
Project Milestones

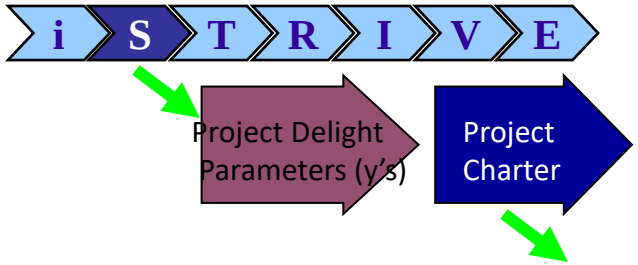
- A detailed plan of the project
- Key steps and dates

Green Belt Program



Project Plan Steps





SET

Business Case	Problem & Goal Statement	Scope	Project Milestones	**Resources; Team Roles & Respblts.**

Resources:

- Roles involved in the project
- Responsibilities of specific members

Get Management Approval after completing the Project Charter

iSTRIVE Project Charter

Project Information

Project Title
Project ID
Project Description

Key Milestones Review

Identify
Set
Track
Review

Improve
Validate
Execute

Process Information

Business Function: Satyam Computers
Product/Service
Primary Potential Areas

Secondary Potential Areas

Project Delight Parameters

Team

Project Sponsor
Process Owner
Team Leader
Quality Project facilitator

Team Members

Business Case (Expected Benefits)

Problem Statement

Project Objective

Project Scope

Project Plan

Resources

Project A,I,C:
Customer:
Investors:
Associates:

Concurrence

Process Owner	Team Leader	Black Belt
Project Sponsor	Functional Manager	Master Black Belt

SET

Questions to be answered

1. What are the Stakeholders' Delight requirements?
2. What are the Project Delight Parameters (Project y's) ?
3. How many Projects are selected to improve the process PI (Y) ?
4. Is the Business Problem defined ?
5. Are the Projects focused (Scope) ?
6. Is the Need for Improvement /Benefits identified ?
7. What are the Team Roles & Responsibilities and Training needs?
8. Do we have Buy-in from all parties & the Management ?

Deliverables - Phase level

1. Stakeholders Delight requirements
2. Project Delight Parameter (y)
3. Business Problem Definition
4. Project Charter

Tools

Tools:		
Voice of A-I-C-S	Affinity Diagram	Bench-marking
Kano Analysis	Theme Prioritization Matrix	QFD
Gantt Chart	Project Charter	COQ Analysis

SET

Do's

1. Identify the stakeholders and do a stakeholder analysis for the process

2. Think of the Set step as symptoms analysis - what does the problem "look like"

3. Define the PDPs (Project y's) in line with A-I-C-S perspective. List the scope & boundaries

4. Consider using more than one PDP per Project. These should also be measured to assure that the chosen solution does not affect the other aspects of the process

5. Assure that resource requirements are identified and available to the project

Don'ts

1. Ignore the rigor of Stakeholder Analysis and arriving at the Measures with the specification. This may be needed to Set the Key Projects leading to significant Process improvements

2. Decide at Sub-process level identification Qualitatively.

3. Proceed without Management approval of the Project(s)

SET

Society Delight Parameters

Associate Delight Parameters

Investor Delight Parameters

Customer Delight Parameters

XXXX
YYYY
ZZZZZ

Voice of Stakeholders

Project Delight Parameters (y's)

Project Objective & Scope

Project Plan

Project Benefits
(Need for Improvement)

Project(s) Approval

Set the Project Delight Parameters (y's) & Goals

SET

From **Identify**

Capture Voice of A-I-C-S

- Questionaire
- Interviews
- Brainstorming
- Group Meetings
- Surveys

Translate Voice of A-I-C-S into Delight Requirements

- Kano Analysis
- Affinity Diagram
- QFD

Analyze Delight requirements wrto Process Deliverables & Circle Measures - Ys

- Stratification
- Prioritization Matrix

Prioritize the key Delight requirements – y's

- PDP Flow

Initiate a Project Charter for each Process measure – y

- Project Charter

Select Project Leader & BB

Set Goals / Targets for PDPs (y's) based on Benchmarks, A-I-C-S expectations, Self determined stretch targets

- Bench-marking
- Update Project Charter

Benchmark PDPs

- Benchmarking

Set the Specifications For PDPs

Scope the Project(s) & Boundaries (In & Out)

- PDP Flow

Set Project Delight Parameters (Project–y's)

- Stratification
- Requirements Analysis (matrix)

Define Business Problem(s)

- Problem Statement
- Data Collection Plan

Select Project Team & Resources

Document the Need for Improvement & Impact

- Background
- Opportunities
- Impact

Document Potential Gains/Opportunities and Quantify Financial terms (Hard & Soft)

- COQ Analysis

Set Project Activities & Timelines

- Gantt Chart
- Project Plan
- Data Collection Plan

Management/ 6σ MBB Review

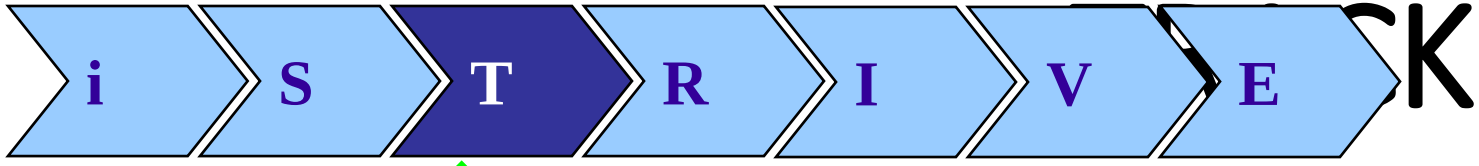
No

Yes

Go to **Track**

TRACK

Module 4



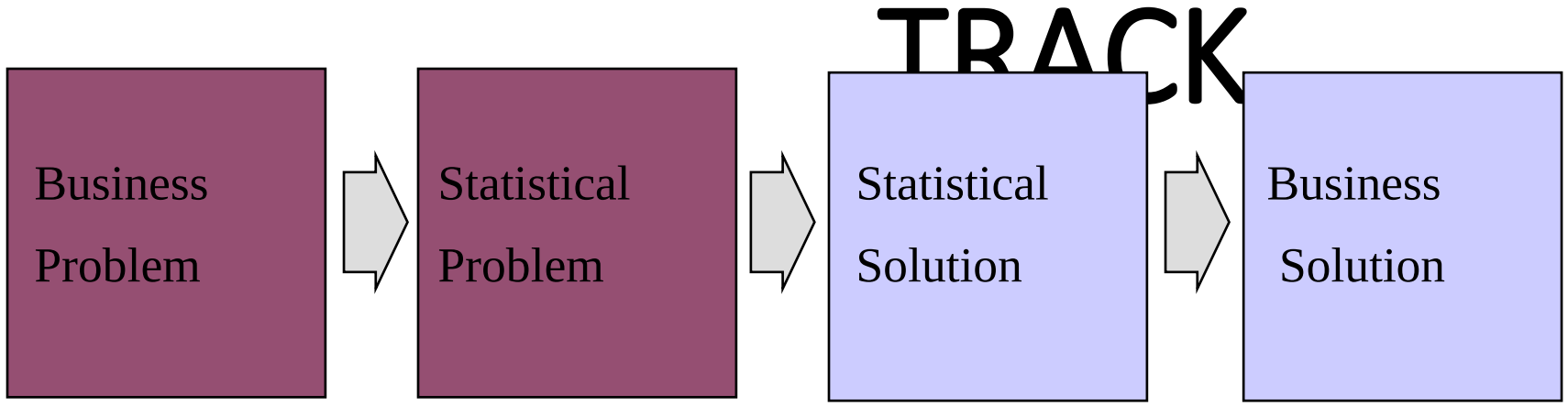
Measurement System / Review Process

- Study the Measurement System of y
- Review R&R to validate the Measurement System
- Improve R&R if variation is high

Baseline Performance and Review Targets

- Collect Data on the performance of y
- Analyze data on y for Stability
- Evaluate the Capability of y (Calculate Present Z value for y)
- Review Baseline Performance with set Targets
- Revise Targets if required
- Update Project Charter and Obtain Management Approval

Track Current Process Capabilities



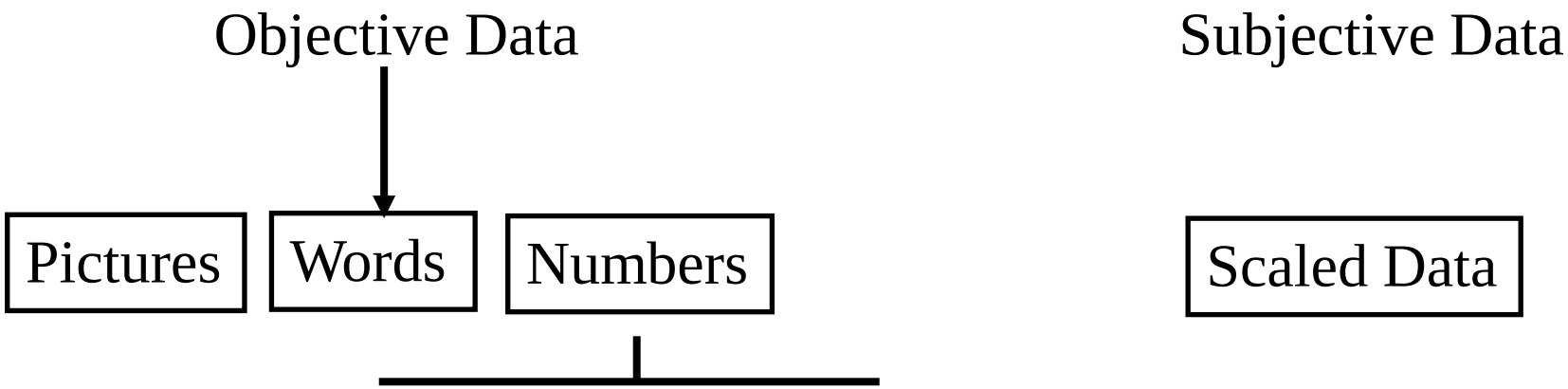
Transform Business Problem to Statistical Problem

Transformation Requires:

- **Measurement of Current Performance**
  - **Gathering Data**
  - **Analyzing Data**

TRACK

Data


Objective Data
Subjective Data
Pictures
Words
Numbers
Scaled Data
Attribute
Variables
Rankings
or ratings

Green Belt Program

Objective Data

Attribute / Discrete Data

- **Pass/Fail** or **Yes/No** decisions
- **No. defective**
- **No. on-time**
- **Blood Group O, A, B, AB**
- **Performance A, B, C & D**

Variable / Continuous Data

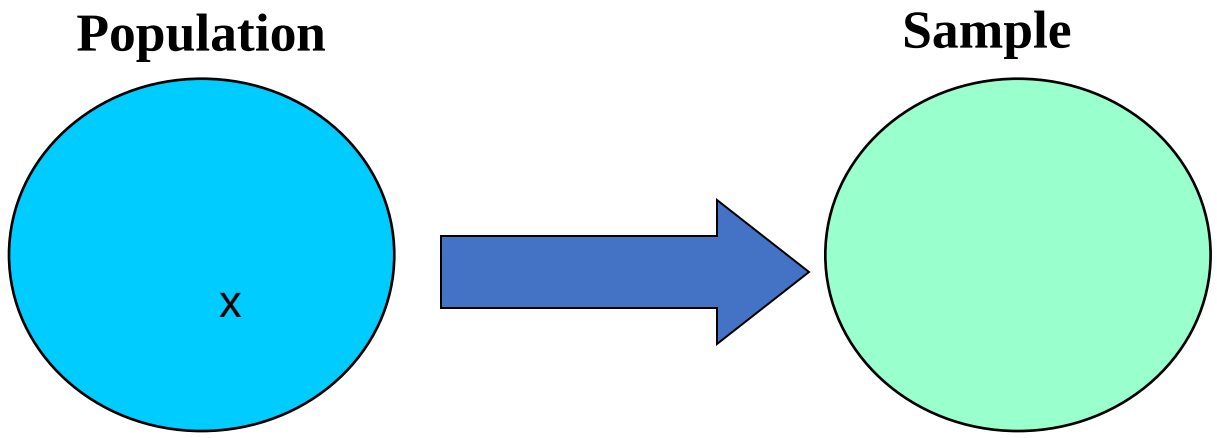
- **Delivery Time**
- **Effort expended**
- **Memory utilization**
- **Cost of rework**
- **Cycle time in hours, days, weeks, etc**

Variable data can be converted to attribute data

TRACK

Sampling is the process of:

Collecting only a portion of the data that is available or could be available, and drawing conclusions about the total population (statistical inference)



N = 50,000

n = 100

Population is 50,000 movie goers of New Delhi

From the sample, we infer that the average time in the ticket queue (y) is 22 minutes

- **Population (N): The entire set of objects or activities for a process**
  - μ : the mean (arithmetic average) calculated for a population
  - σ : the standard deviation calculated for a population

- **Sample (n): a group that is a part or subset of a population**
  - x : the mean (arithmetic average) of a sample
  - s : the standard deviation of a sample

TRACK

Project Location		Name Dates			
Testing Information		**Module 1**	**Module 2**	**Module 3**	**Module 4**
Software Defect					
	Requirements				
	Design				
	Coding				
	Operational document				
	Test Case				
Other Defects					
	Hardware				
	Operating System				
	User Mistake				
	New requirement				

TRACK

Characteristics of Data

Measures of Central Tendency (Location)

1. Mean
2. Median
3. Mode

Measures of Spread/ Dispersion (Variation)

1. Range
2. Standard Deviation

Mean = The average value (the arithmetic average)

$$\overline{X} \quad \frac{\text{Sum of the data points}}{\text{Number of data points}}$$

Sl. No.	1	2	3	4	5	6	7	8	9	10	11
Obsevations	12	14	11	12	12	12	15	17	22	15	12

- Sum of data points = 154
- Number of data points = 11
- Mean = 154/11 = 14

Mode = The data point having the highest frequency (Maximum Occurrences)

Sl. No.	1	2	3	4	5	6	7	8	9	10	11
Obsevations	12	14	11	12	12	12	15	17	22	15	12

Mode = 12, as 12 occurs 5 times in the data set shown.

Median = the 50% point, or the "middle number"

TRACK

To find the median of a data set:

(1) arrange data in order from smallest to largest

(2) the middle number is the median!

Sl. No.	1	2	3	4	5	6	7	8	9	10	11
Obsevations	12	14	11	12	12	12	15	17	22	15	12

Arranged data

Sl. No.	1	2	3	4	5	6	7	8	9	10	11
Obsevations	11	12	12	12	12	12	14	15	15	17	22

Median = 12 in the data set shown

Variation TRACK

Assignable (Special Cause)

- Sporadic
- Very Few
- Large Effect
- Easy to identify and eliminate

Chance (Random Cause)

- Random
- Large in number
- Small Effect
- Difficult to identify and eliminate

85 % of the total variation is caused by Chance Causes

To describe how far the data is spread

Range = R; the difference between largest and smallest observations (minimum & maximum data points)

Moving Range = MR; the difference between consecutive observations

Standard Deviation = s

Variance = s^2 (just the square of the std dev!)

Range

Sl. No.	1	2	3	4	5
Obsevations (x)	12	13	11	12	12

Range in the data set shown above is 13 -11 = 2

Variance and Standard Deviation

Step 1: Find the mean of the data (x = 12)

Step 2: Find the difference of each data point from the mean

Sl. No.	1	2	3	4	5
Obsevations (x)	12	13	11	12	12
x-mean	0	1	-1	0	0

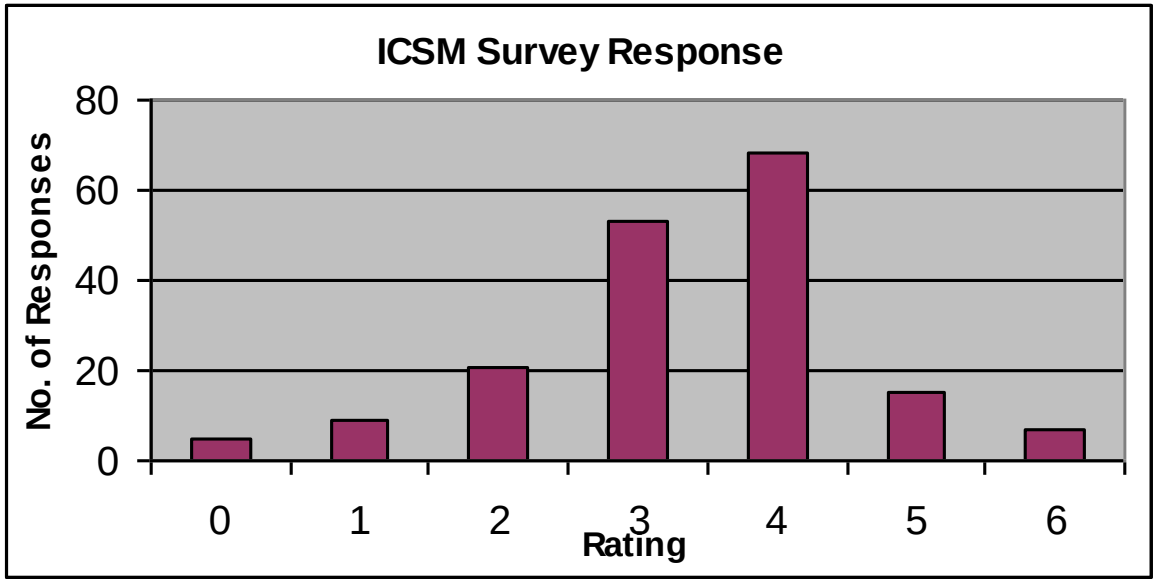
Step 3: Square the differences and add = 1+ 1 = 2

Step 4: Variance = Divide by (n-1) = 2/4 = 0.50

Step 5: Standard Deviation = Take Square root = √ 0.50 = 0.7071

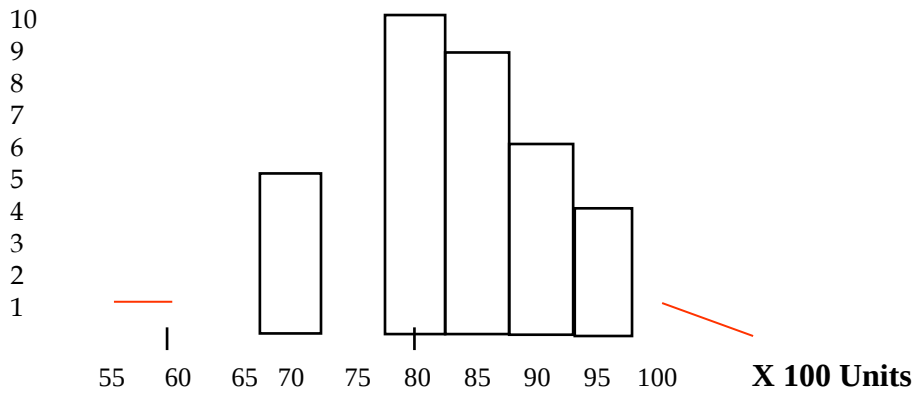
Discrete Data : Survey Response

ICSM Survey Responses for a particular Circle	
Rating	**No. of Responses**
0	5
1	9
2	21
3	53
4	68
5	15
6	7
Total	**178**



Power Consumption in one location for 48 days

Power Consumption (x100 Units)	No. of Days
55	1
60	1
65	4
70	5
75	7
80	10
85	9
90	6
95	4
100	1



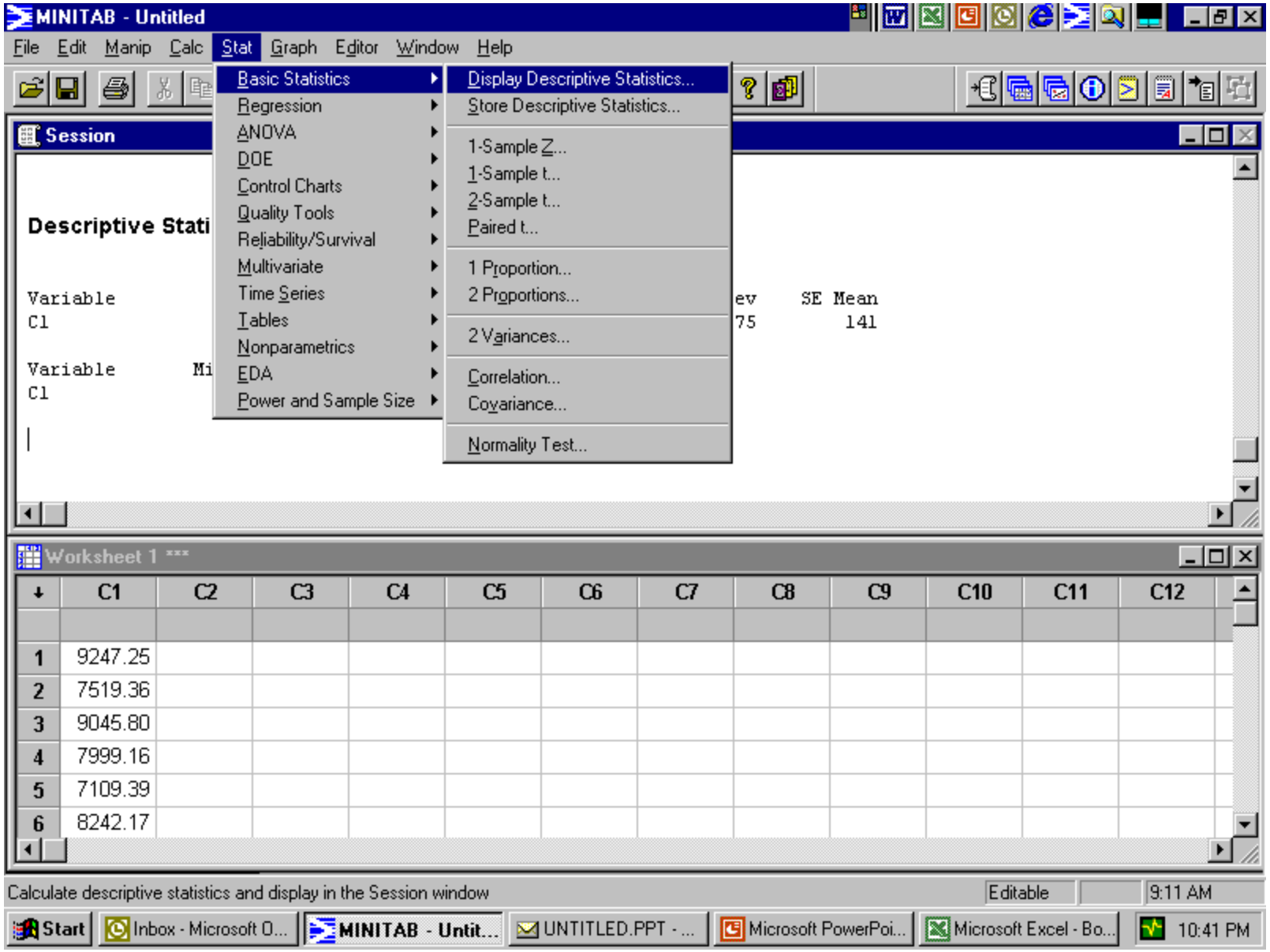
Power Consumption	
Day	Units
1	9247.25
2	7519.36
3	9045.8
4	7999.16
5	7109.39
6	8242.17
7	7677.22
8	8573.72
9	7839.44
10	7852.62
11	7996.33
12	8094.72
13	7096.52
14	8414.7
15	8454.36
16	7529.77
17	743101
18	7988.63
19	6450.53
20	8329.32
21	7766.73
22	9468.25
23	8643.7
24	6894.45
25	9246.41
26	891134
27	7932.33
28	8038.31
29	9769.72
30	8227.39

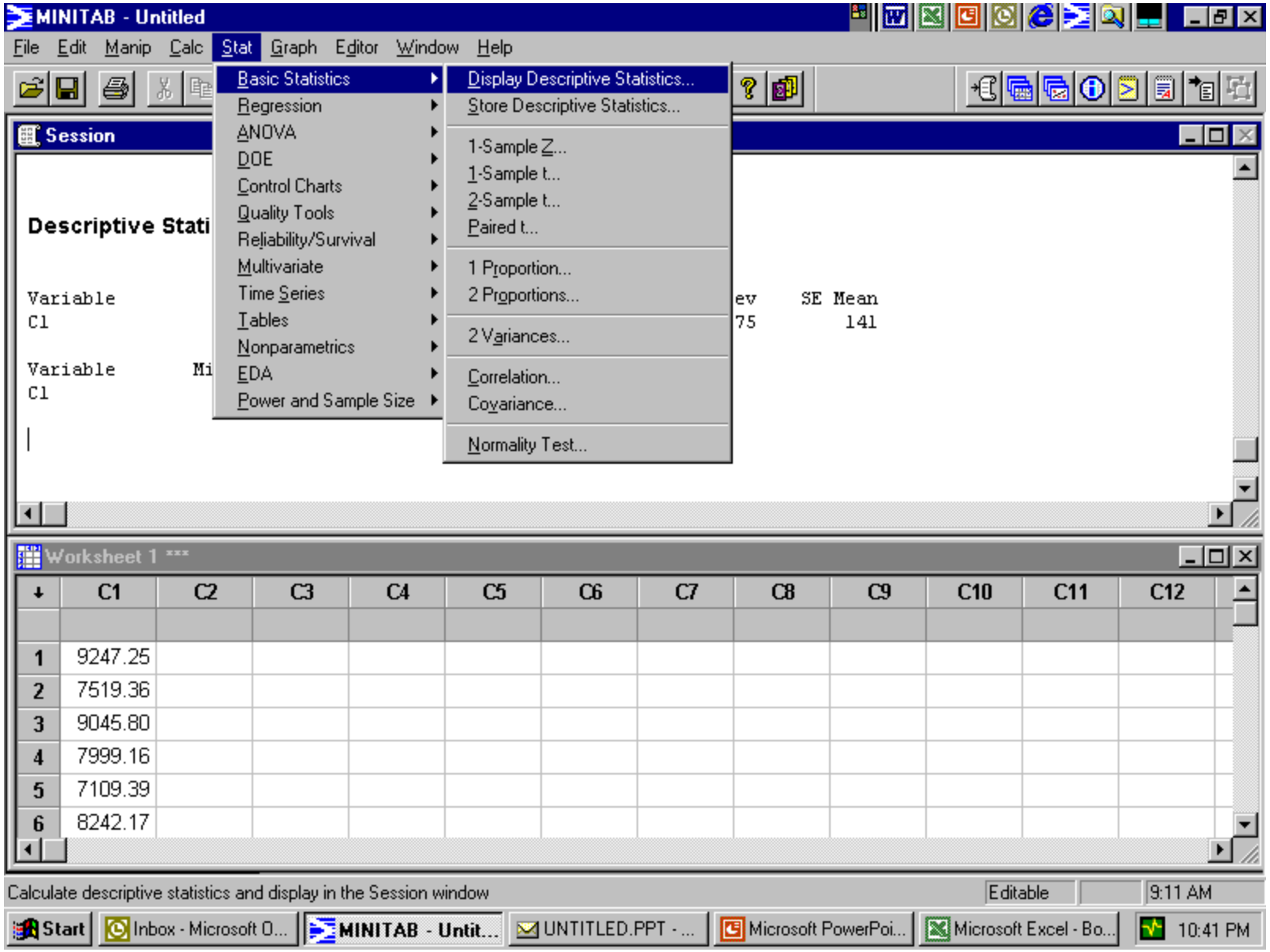
No. of data points	30
Mean (x-bar)	8126.4
Std. Deviation (s)	774.9
Maximum	9769.7
Minimum	6450.5
Range	3319.19

When

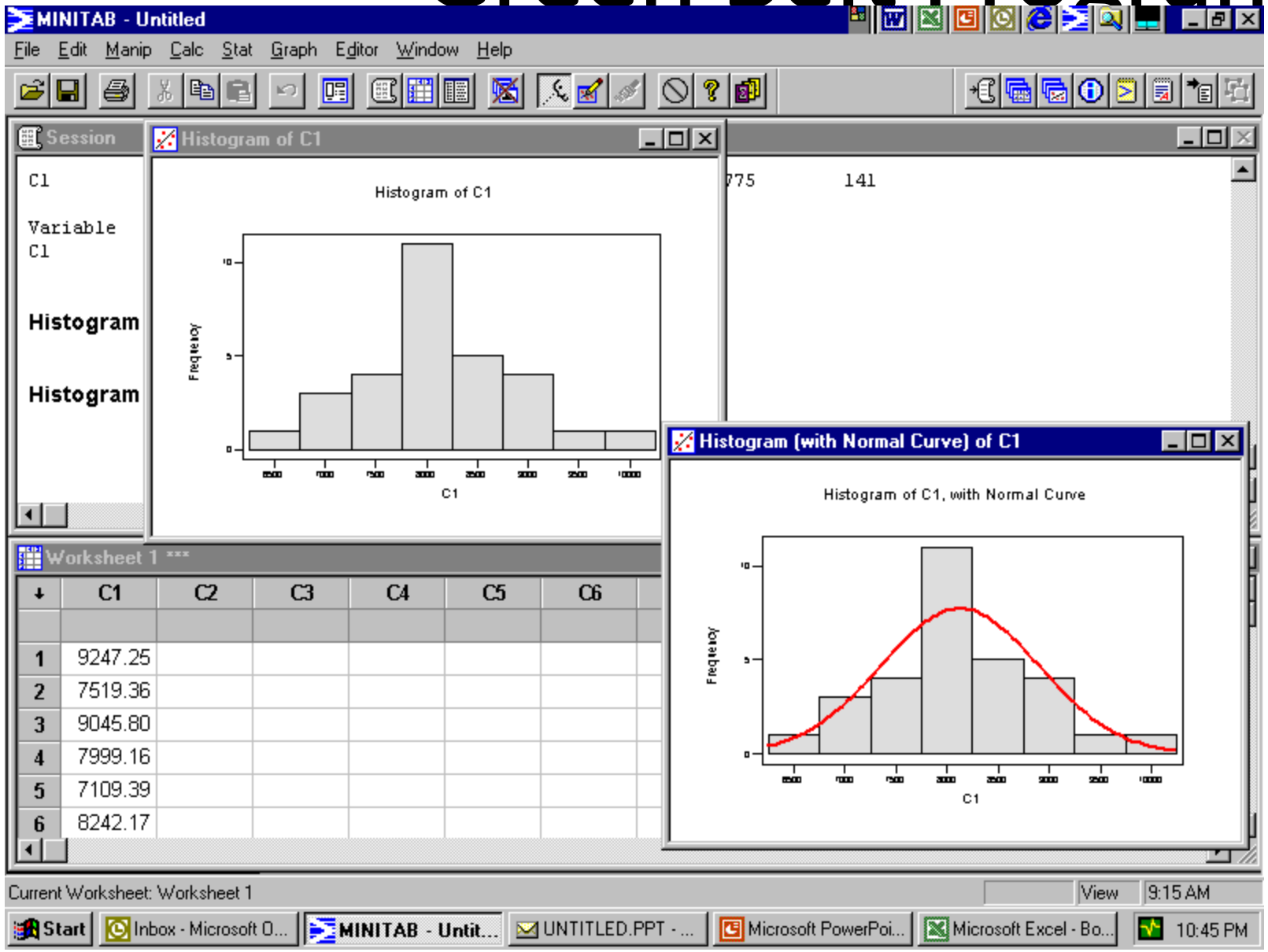
• To determine if process is on target meeting customer requirements.

• To determine if variation in process is normal or if something has caused it to vary in an unusual way

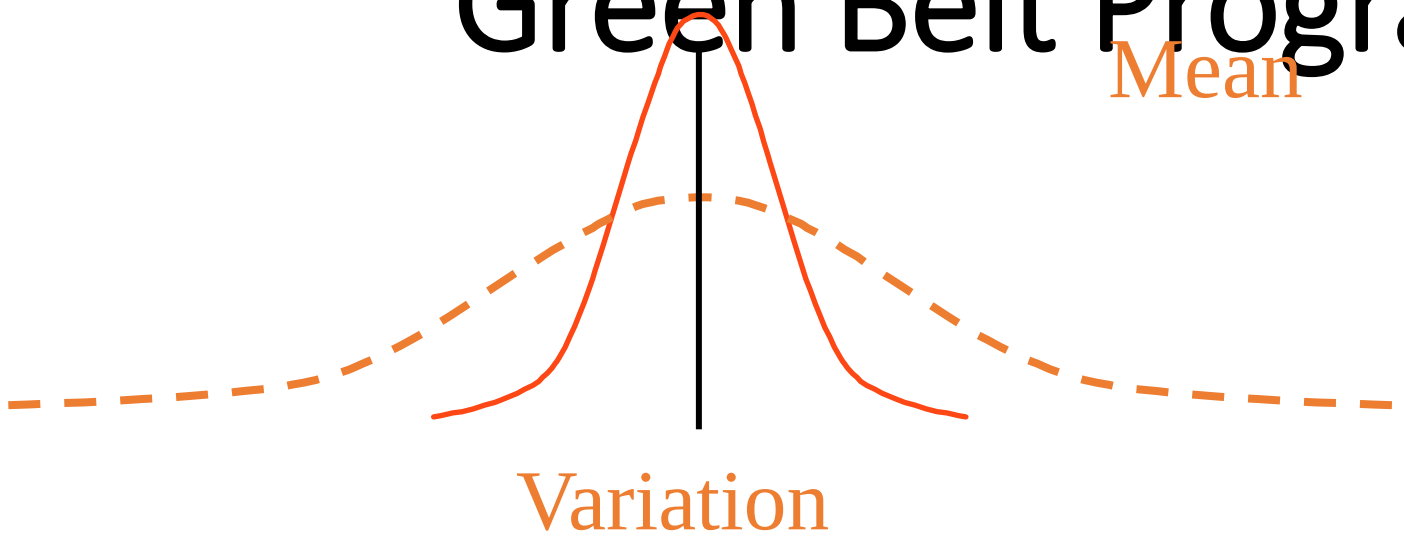

MINITAB - Untitled
File Edit Manip Calc Stat Graph Editor Window Help
Basic Statistics
Regression
ANOVA
DOE
Control Charts
Quality Tools
Reliability/Survival
Multivariate
Time Series
Tables
Nonparametrics
EDA
Power and Sample Size
Display Descriptive Statistics...
Store Descriptive Statistics...
1-Sample Z...
1-Sample t...
2-Sample t...
Paired t...
1 Proportion...
2 Proportions...
2 Variances...
Correlation...
Covariance...
Normality Test...
Session
Descriptive Stati
Variable
C1
ev SE Mean
75 141
Variable Mi
C1
Worksheet 1 ***
C1 C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12
1 9247.25
2 7519.36
3 9045.80
4 7999.16
5 7109.39
6 8242.17
Calculate descriptive statistics and display in the Session window
Editable
9:11 AM
Start
Inbox - Microsoft O...
MINITAB - Untit...
UNTITLED.PPT - ...
Microsoft PowerPoi...
Microsoft Excel - Bo...
10:41 PM


MINITAB - Untitled
File Edit Manip Calc Stat Graph Editor Window Help
Basic Statistics
Regression
ANOVA
DOE
Control Charts
Quality Tools
Reliability/Survival
Multivariate
Time Series
Tables
Nonparametrics
EDA
Power and Sample Size
Display Descriptive Statistics...
Store Descriptive Statistics...
1-Sample Z...
1-Sample t...
2-Sample t...
Paired t...
1 Proportion...
2 Proportions...
2 Variances...
Correlation...
Covariance...
Normality Test...
Session
Descriptive Stati
Variable
C1
ev SE Mean
75 141
Variable Mi
C1
Worksheet 1 ***
C1 C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12
1 9247.25
2 7519.36
3 9045.80
4 7999.16
5 7109.39
6 8242.17
Calculate descriptive statistics and display in the Session window
Editable
9:11 AM
Start
Inbox - Microsoft O...
MINITAB - Untit...
UNTITLED.PPT - ...
Microsoft PowerPoi...
Microsoft Excel - Bo...
10:41 PM

Green Belt Program



Green Belt Program



Mean of Curve ‘1’ is more representative of its data set as compared to Curves ‘2’ & ‘3’

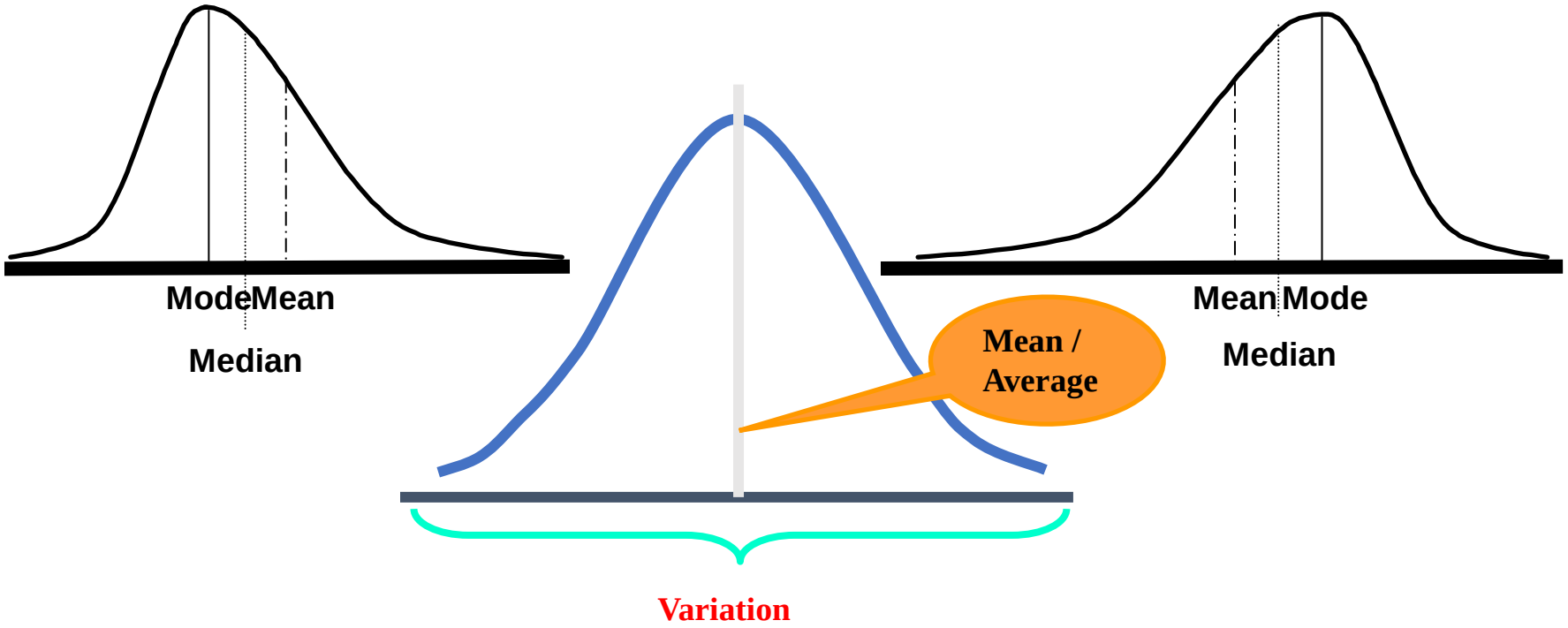
Spread outside the specifications may result in defects; this information is not provided by mean

Symmetric Data set

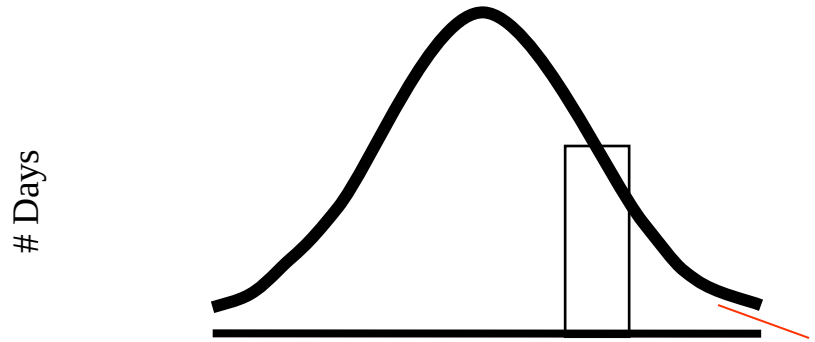
- A data set in which spread of the data set around its mean is identical
- For such a data set - **mean = mode = median** **< > Normal Distribution**

Asymmetric Data set

- Positive / Right skewed - high spread on the right side of the mean
- Negative / Left skewed - high spread on the left side of the mean



Power Consumption in one location for 48 days



- Frequency distributions made from sample data.
- The sample process data characteristics (average, standard deviation, shape, etc.) are studied to predict the future behavior by fitting an appropriate probability distribution.

Green Belt Program

Variables Data	Discrete Data
Normal Distribution	Binomial Distribution
Exponential Distribution	**Poisson Distribution**
Weibull Distribution	Hyper-Geometric Distribution
Rayleigh Distribution	Etc.
Beta Distribution	
Etc	

Probability distribution is a ‘theoretical’ frequency distribution

Green Belt Program

Applicable in situations where multiple defects / failures can occur in an item but the probability of each defect is very low.

Characterized by single parameter λ which is called shape parameter and represents the average no. of defects in an item. λ= Total Defects/Total Items = D/U = DPU (Defects per Unit)

The probability of finding **n** defects in an item is represented as

$e^{-\lambda} \lambda^n/n!$, where e=2.71828 and n! is 1x2x3x…..xn

Example: In a bill with average no of defects = 2 (λ =2), the probability of having 0 defects (n=0) is $e^{-2}\ 2^0/0! = e^{-2} = 0.135$

•A probability distribution, where the pattern of measurements follows a Symmetrical Bell Shape.

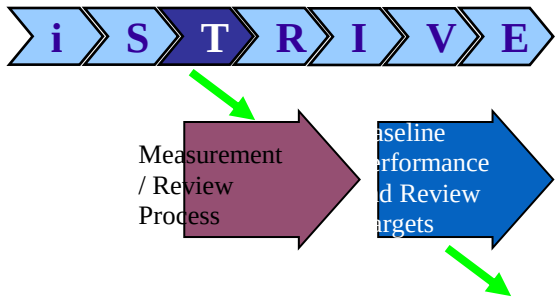
•The curve is tallest in the middle and tails-off symmetrically in both directions.

•Characterized by Mean(μ) and Standard Deviation (σ)

•Most of the Process/Product Characteristics in nature follow Normal Distribution. $y \sim N(\mu; \sigma)$,

where μ = mean and

σ = standard deviation



Performance Baseline - Z

Continuous Data

Location

Spread

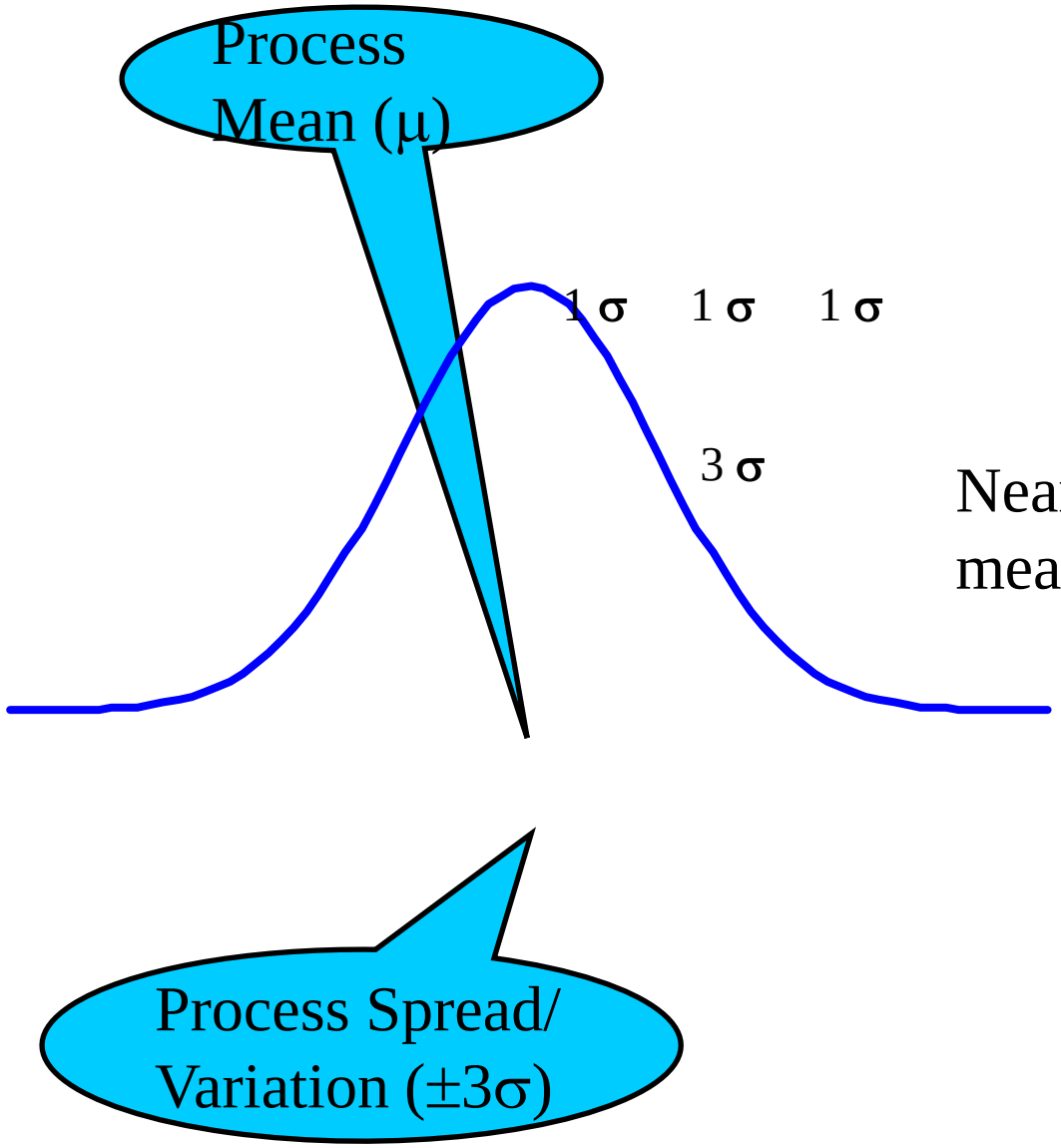
Shape

Discrete Data

Defects

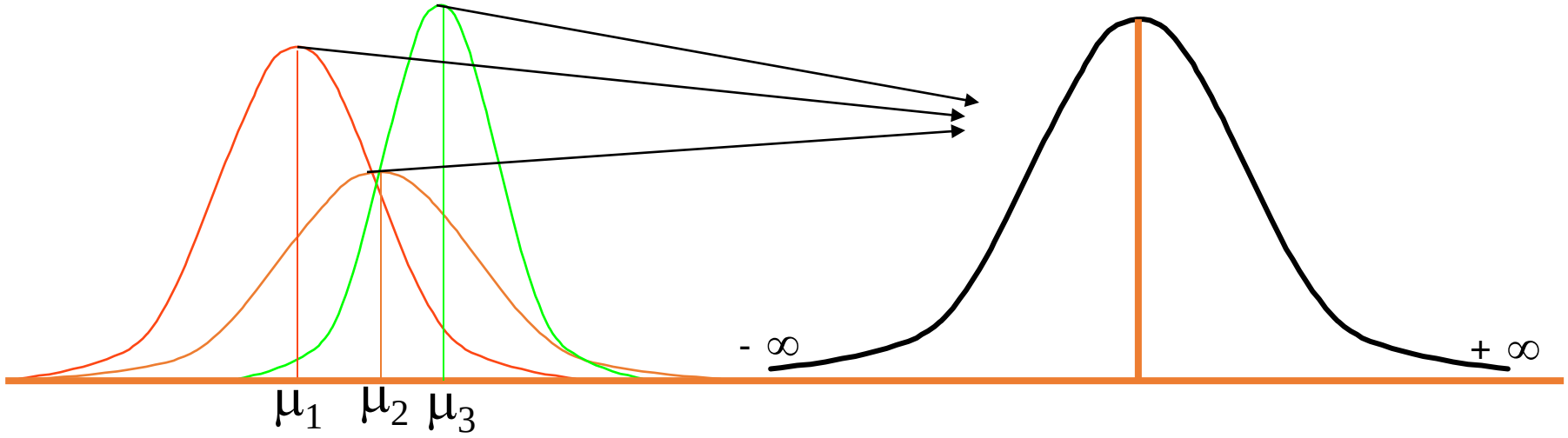
Defects per Unit
(DPU)

Defects per Million Opportunities
(DPMO)



Nearly 100 % of process measurement fall between **+/- 3 σ**

Green Belt Program



Each Normal distribution looks different

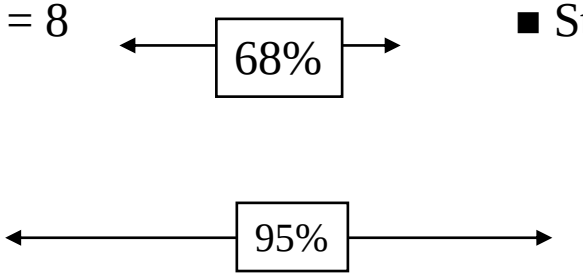
Solution: Z transformation

$$Z = \frac{X-\mu}{\sigma}$$

Green Belt Program

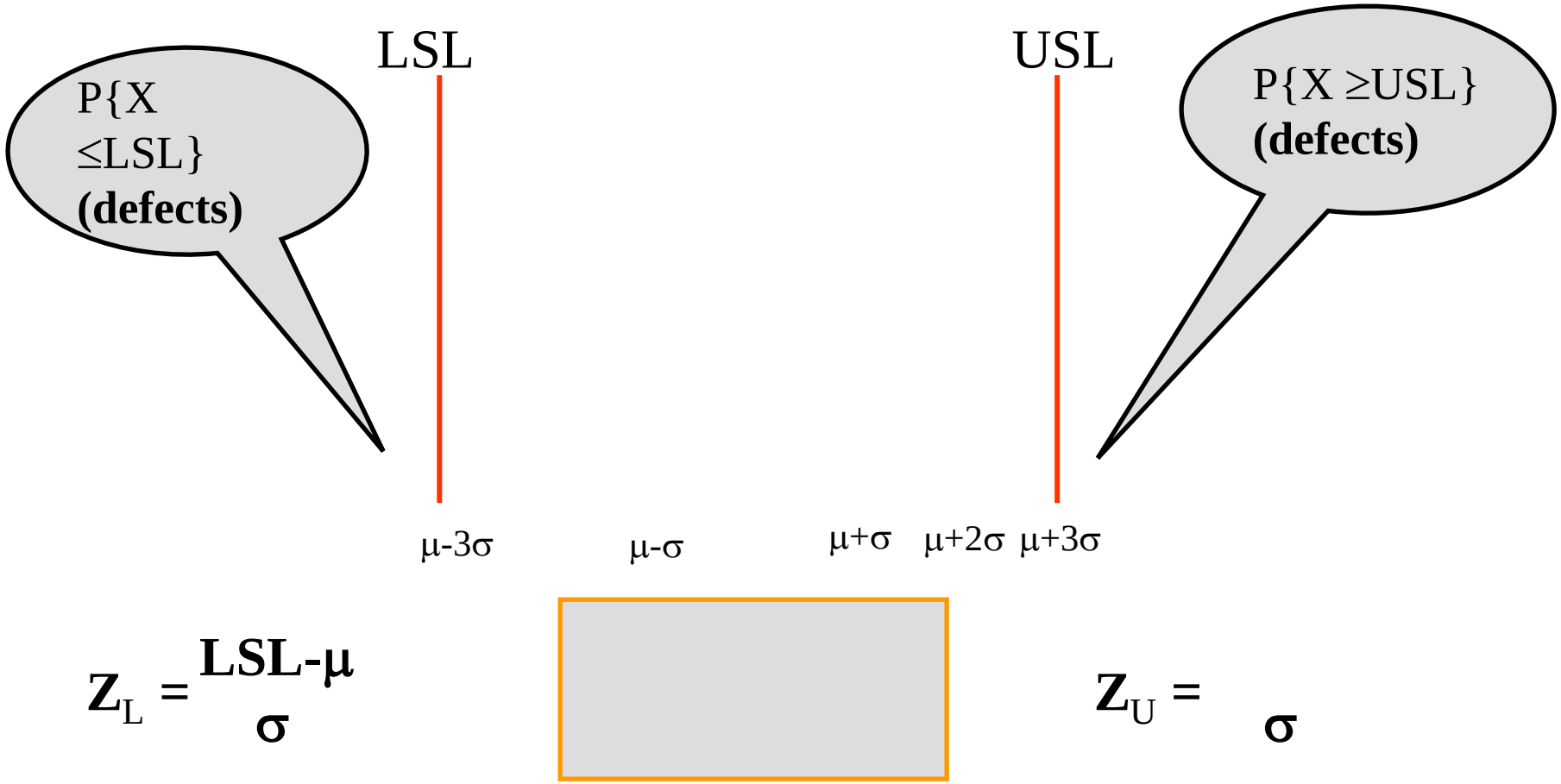
Normal Distribution
Uses Actual Data:
- Average = 40
- Standard Deviation = 8

Standard Normal Distribution
Standardizes Data To:
- Average = 0
- Standard Deviation = 1



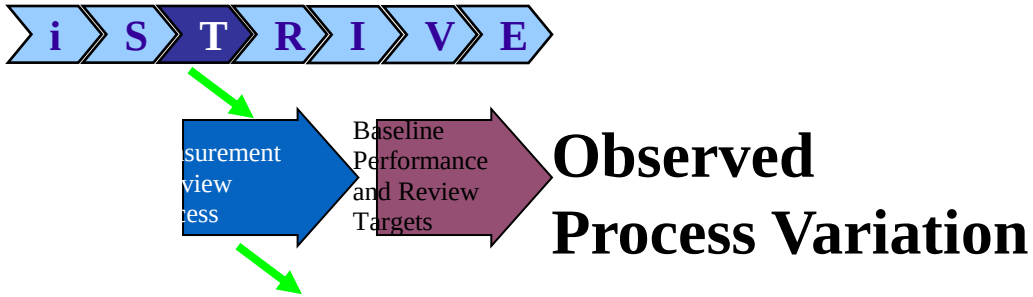
16	24	32	40	48	56	64
-3	-2	-1	0	1	2	3

Z is a unit of measure that is equivalent to the number of standard deviations a value is away from the mean value



$$Z_L = \frac{LSL-\mu}{\sigma}$$

$$Z_U = \frac{}{\sigma}$$



Observed Process Variation

Variation due to Measurement System

Variation due to Process (design & execution)

Variation Due to Measuring System (instrument /Gage)

Variation Due to Appraiser (Associate)

Controllable Variation (Sporadic)

Un-Controllable Variation (Natural)

Accuracy
Stability
Linearity
Repeatability

Reproducibility

Ensure Measurements are repeatable & reproducible

The capability of a process is defined as the inherent variability of a process in the absence of any undesirable special causes and the variability is solely due to common causes.

Short and Long-term Capability Studies

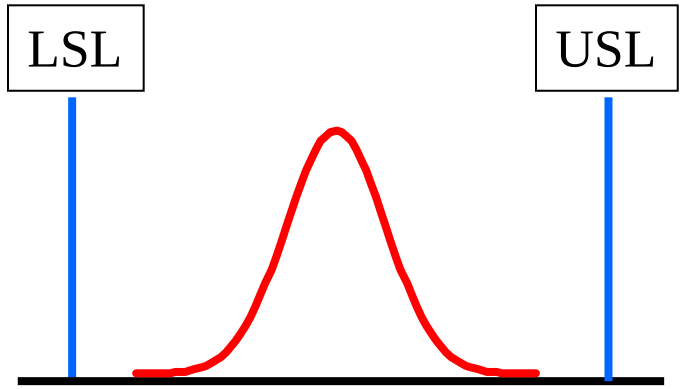
- A Short-term Capability study covers a relatively short period of time (Hours, Days) consisting of 30 to 50 data points

- A Long-term capability study covers a relatively long period of time (Weeks, Months) consisting of 100-200 data points

Process Spread Vs. Specification Spread

Case I: A Highly Capable Process

The Process Spread is Well Within the Specification Spread



$6\sigma < [USL-LSL]$

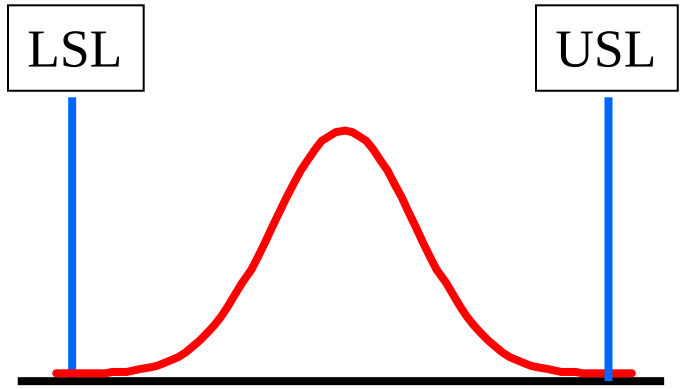
Customers experience less difficulty, less rework, more reliability…Should translate into higher profits.

Green Belt Program

Process Spread vs. Specification Spread

Case II: **A Barely Capable Process**

The Process Spread is just about matches 6 σ



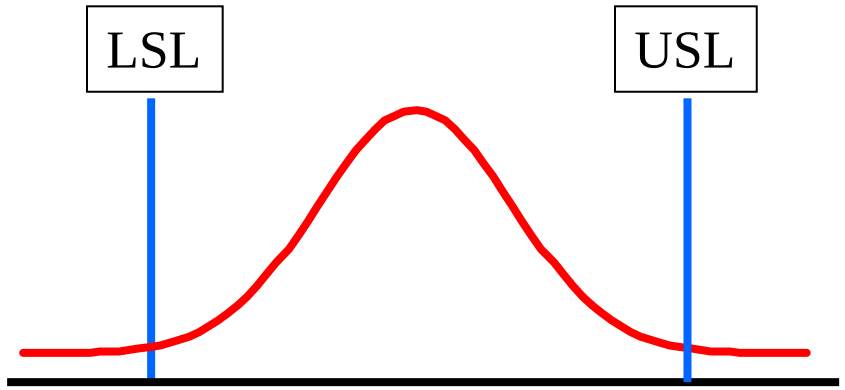
6σ = [USL-LSL]

Process barely capable of meeting specifications…Should watch closely to detect shifts.

Process Spread vs. Specification Spread

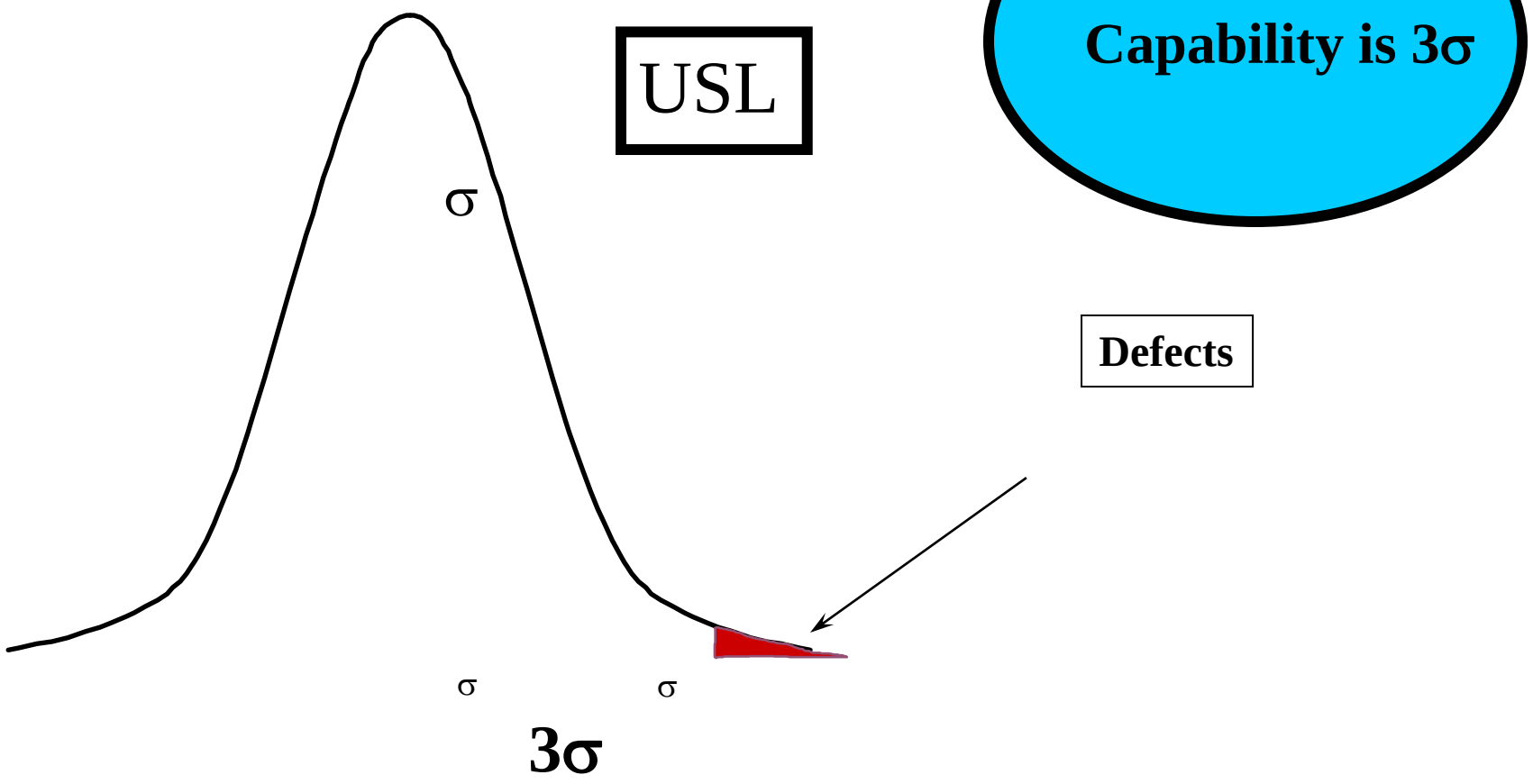
Case III: A Not Capable Process

The Process Spread is beyond 6 σ



6σ > [USL-LSL]

Process not capable of meeting specifications regardless of where the process mean is located...Leads to many defects



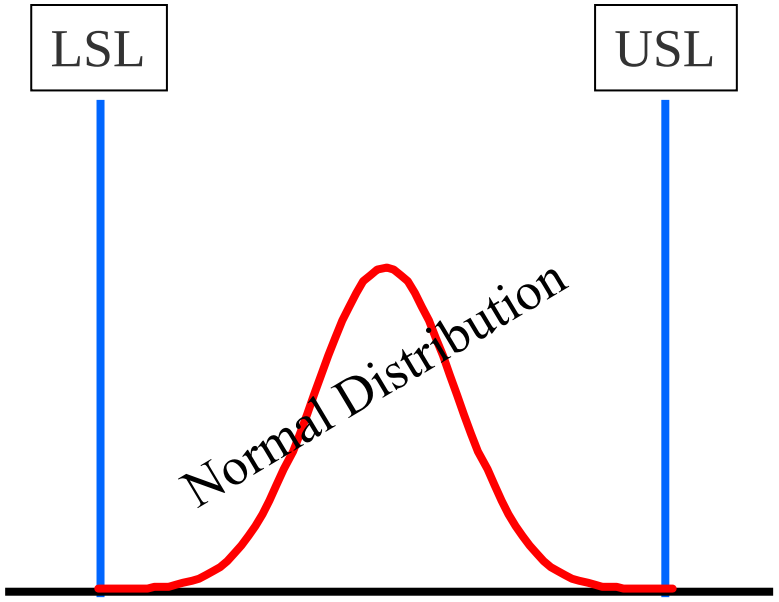
Capability is 3σ

Defects

Process Capability Index: C_p

The simplest capability index, C_p, is the ratio of the specification spread to the process spread, the latter represented by six standard deviations or 6σ.

$$C_p = \frac{(\text{Maximum Allowable Range of Characteristic})}{(\text{Normal Variation of the Process})}$$



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

$$Z = 3C_p$$

The capability index, C_{pk} reflects the current process mean's proximity to either the upper specification limit (USL) or the lower specification limit (LSL).

$$C_{pk} = \text{Min}\left(\frac{\bar{X} - LSL}{3s}, \frac{USL - \bar{X}}{3s}\right)$$

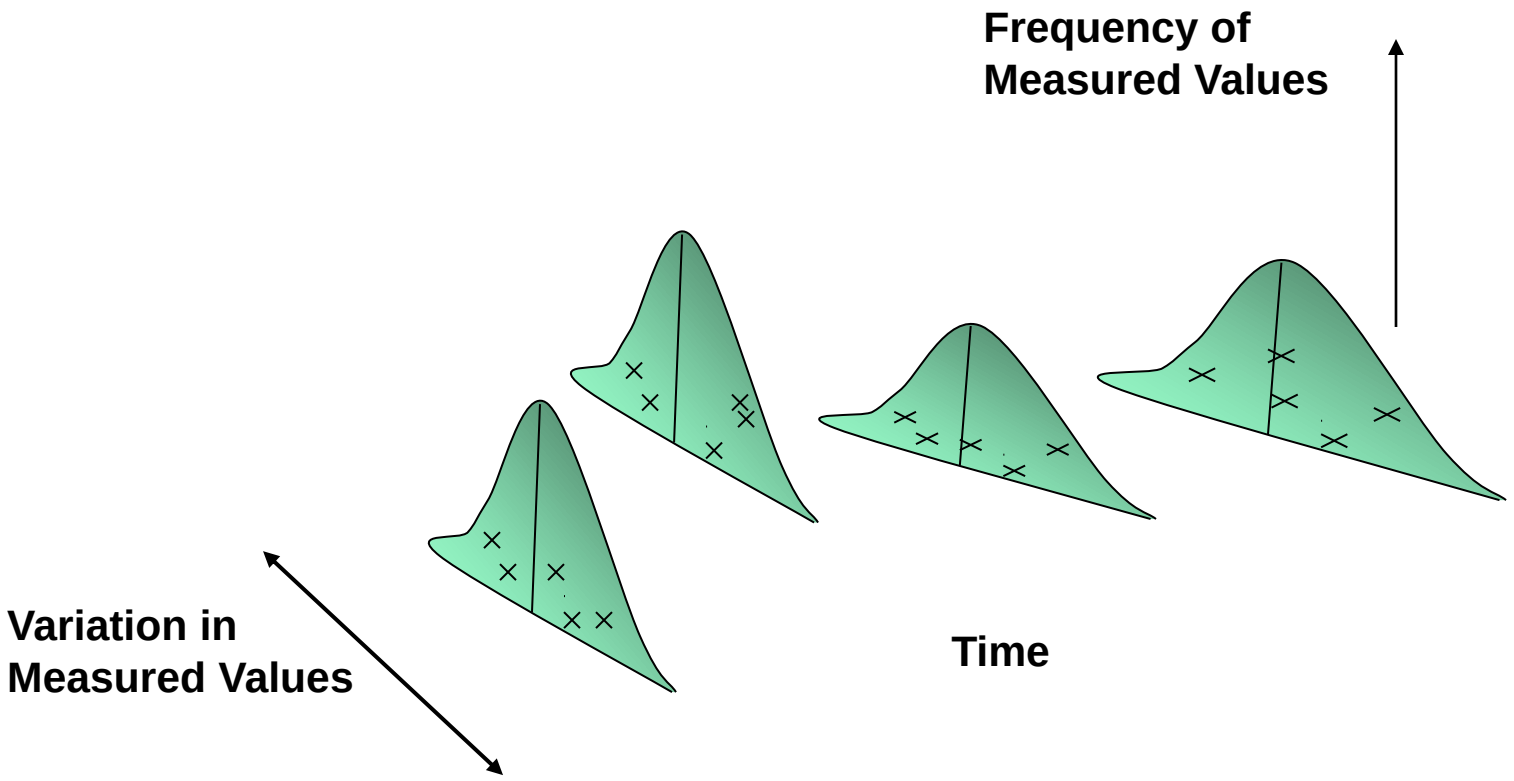
Cpk = Cp(1- k)

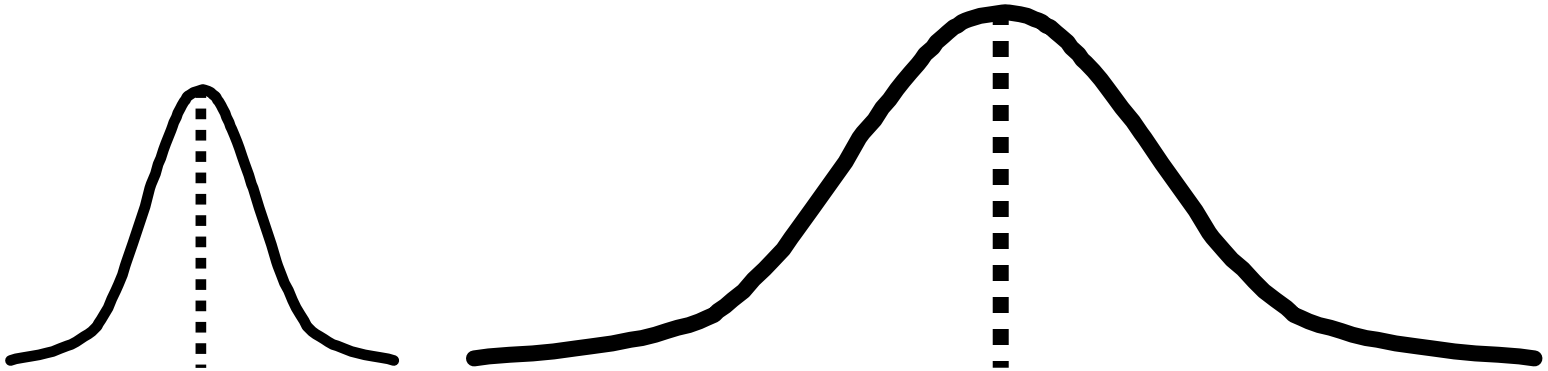
where k =

T - Target Mean

This index accounts for the static mean shift in the process.

Process Shift & Drift



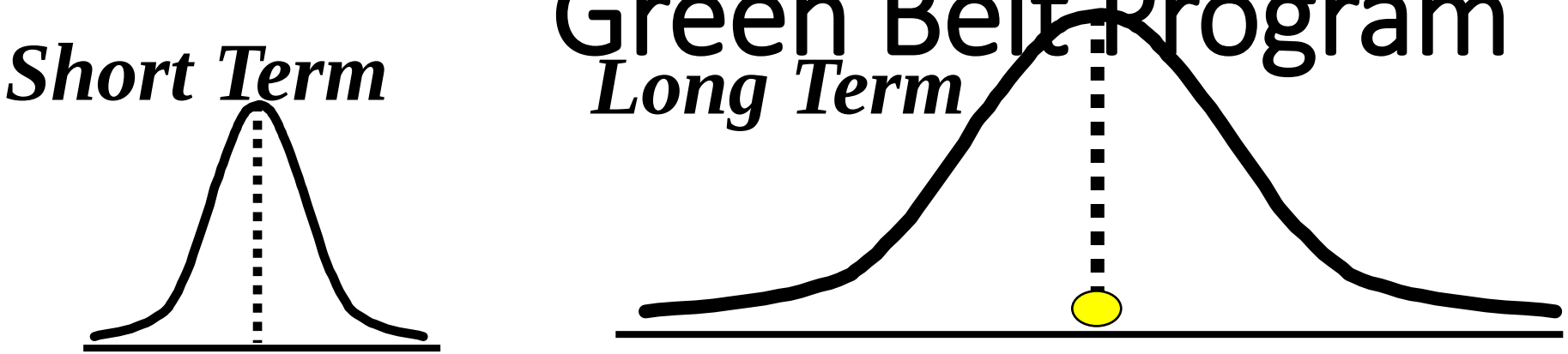


Short Term *Long Term*

$$Z_{ST} = 1.5 + Z_{LT}$$

A Process appears to perform better in the short term

Green Belt Program



smaller sigma

larger Z

larger sigma

smaller Z

1.5

Compute Z_{ST} given LSL and long-term data.

Long-term data: $\sigma_{LT} = 0.0594$ and $\mu_{LT} = 1.052$

Specifications: LSL = 0.9

Compute Z_{LSL} , $p(Defects_{LSL})$, and Z_{ST}

LSL = 0.9

$p(Defects_{LSL}) = 0.0052$

Long-Term Z Distribution

$Z_{LSL} = -2.5589$ $\mu = 1.052$

$Z_{LSL} = Z_{LT} = (0.9-1.052) / 0.0594 = -2.5589$

$p(Defects_{LSL}) = NORMSDIST(Z_{LSL}) = 0.0052$

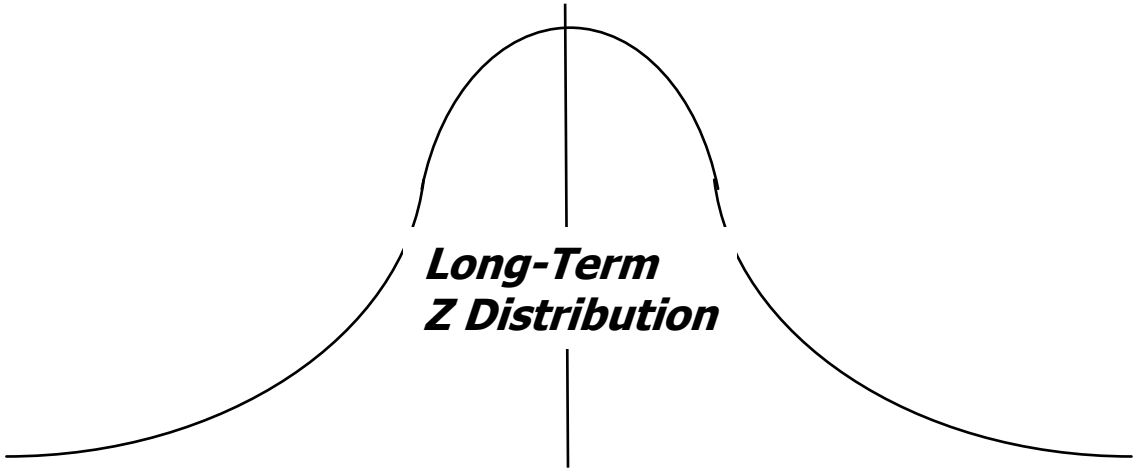
$p(Defects_{LSL}) = 0.0052$ $Z_{LT} = 2.56$ $Z_{ST} = 4.06$

Exercise 2 - Compute Z_{LT} given USL and long-term data.

Long-term data: $\sigma_{LT} = 0.0594$ and $\mu_{LT} = 1.052$

Specifications: USL = 1.1

Compute Z_{USL} , $p(Defects_{USL})$, and Z_{ST}



$Z_{USL} = Z_{LT} =$ (1.1-1.052) / 0.0594 = 0.8081

$p(Defects_{USL})$ = 1 - NORMSDIST(Z_{USL}) = 0.2095

$p(Defects_{USL}) = 0.2095$ $Z_{LT} = 0.81$ $Z_{ST} = 2.31$

Exercise 3 - Compute Z_{ST} given LSL and USL and long-term data.

Long-term data: σ_{LT} = 0.0594 and μ_{LT} = 1.052
Specifications: USL = 1.1 and LSL = 0.9
Compute Z_{USL} , Z_{LSL} , p(Defects), and Z_{ST}

P(Defects) From Exercises 1 & 2:

LSL = 0.9 **USL = 1.1**

$p(Defects_{LSL}) = 0.0052$

$p(Defects_{USL}) = 0.2095$

Long-Term Z Distribution

Z_{LSL}= - 2.5589 μ = 1.052 Z_{USL}= + 0.8081

$P(Defects) = p(Defects_{LSL}) + p(Defects_{USL}) = 0.0052 + 0.2095 = 0.2147$

Z_{LT} = NORMSINV(1- p(Defects)) = 0.7896

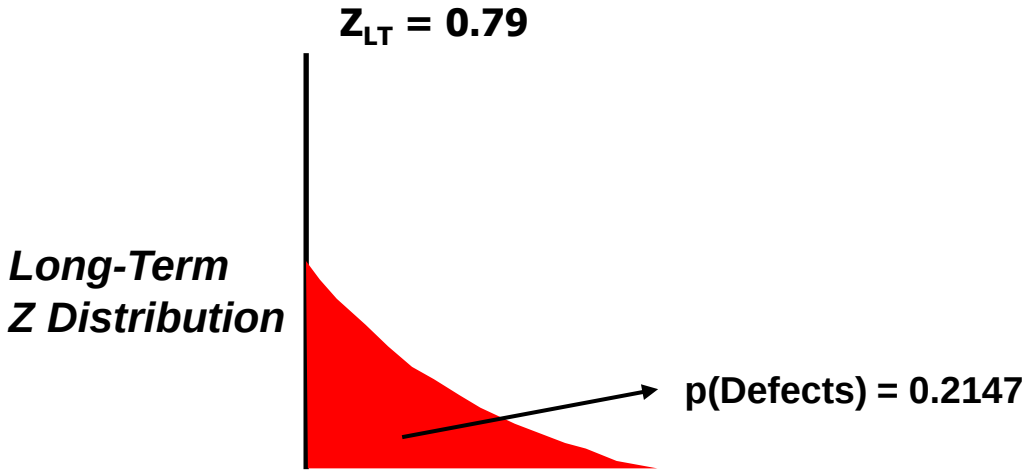
P(Defects)= 0.2147 Z_{LT} = 0.79 Z_{ST} = 2.29

Green Belt Program

Long-term data: σ_{LT} = 0.0594 and μ_{LT} = 1.052

Specifications: USL = 1.1 and LSL = 0.9

Compute p(Defects), and Z_{ST}



P(Defects) = p(Defects$_{LSL}$) + p(Defects$_{USL}$) = 0.0052 + 0.2095 = 0.2147

Z_{LT} = NORMSINV(1- p(Defects)) = 0.7896

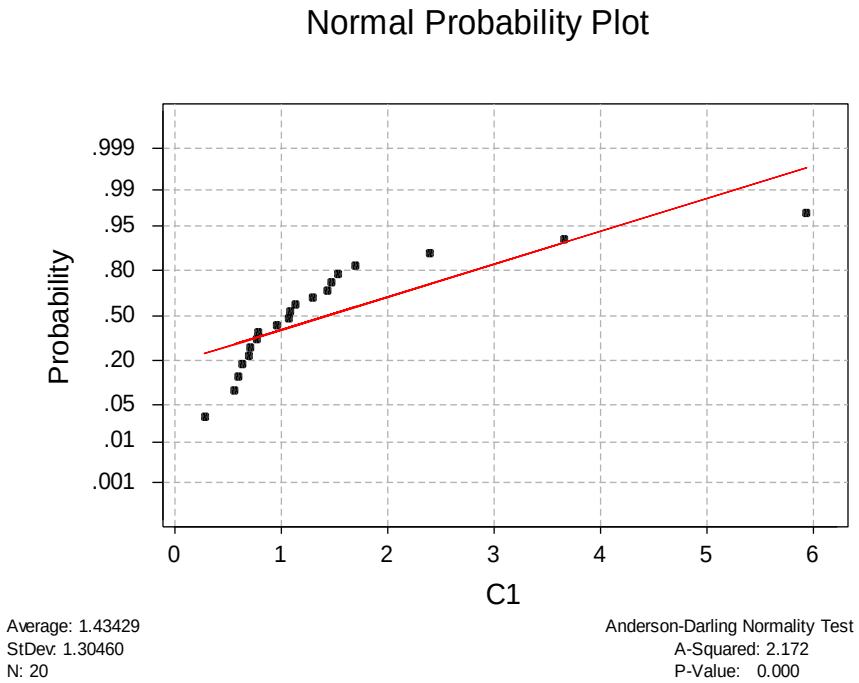
P(Defects)= 0.2147 Z_{LT} = 0.79 Z_{ST} = 2.29

Green Belt Program

All this while, we have assumed the population to be normally distributed

Minitab fits data into a normal curve by default

We must test the normality assumption before doing the capability analysis



Green Belt Program

What if the P-value is less than 0.05?
In such cases, we transform the data points & specifications to convert data into a normal data
Some popular transformations
Y^2
Y^3
$Y^{0.5}$
$1 / Y$
$1 / Y^{0.5}$
LN (Y)
LOG(Y)
Transformation used for Y is also applied to LSL & USL

Use transformed data throughout the analysis and after analysis, reverse transform the results

Green Belt Program

The average values (x) of different samples of size n obtained from any population tend to be normally distributed.

The mean of these average values (x) approaches the population mean μ

The standard deviation of these average values approaches ($\sigma/\sqrt{n}$), where σ is the population standard deviation and n is the sample size

$$\mathbf{x \sim N(\mu, \sigma/\sqrt{n})}$$

x

x_1 x_2 x_n

Green Belt Program

Business criteria to select a sample size include cost, time & effort

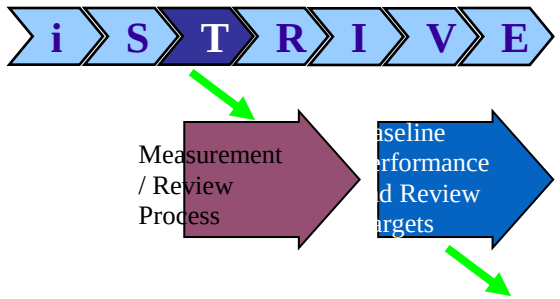
Statistical criteria include the accuracy of the sample representing the population

Higher the sample size, better the accuracy of the information (μ & σ)

There must be a balance between the business & statistical criteria

Generally a sample of size $\geq$30 is sufficient for drawing conclusions about the population

Green Belt Program



Performance Baseline - Z

Continuous Data

- **Location**
- **Spread**
- **Shape**

Discrete Data

- **Defects**
- **Defects per Unit** (DPU)
- **Defects per Million Opportunities** (DPMO)

Green Belt Program

Unit (U)

A tangible and measurable characteristic of a process input/output. The number of parts, sub-assemblies, assemblies, or systems ***inspected or tested.*** Ex:

- In a ticket booking counter, each ticket booked could be a Unit.
- In a pizza delivery case, each order could be a Unit
- In a call center, every customer call could be a Unit

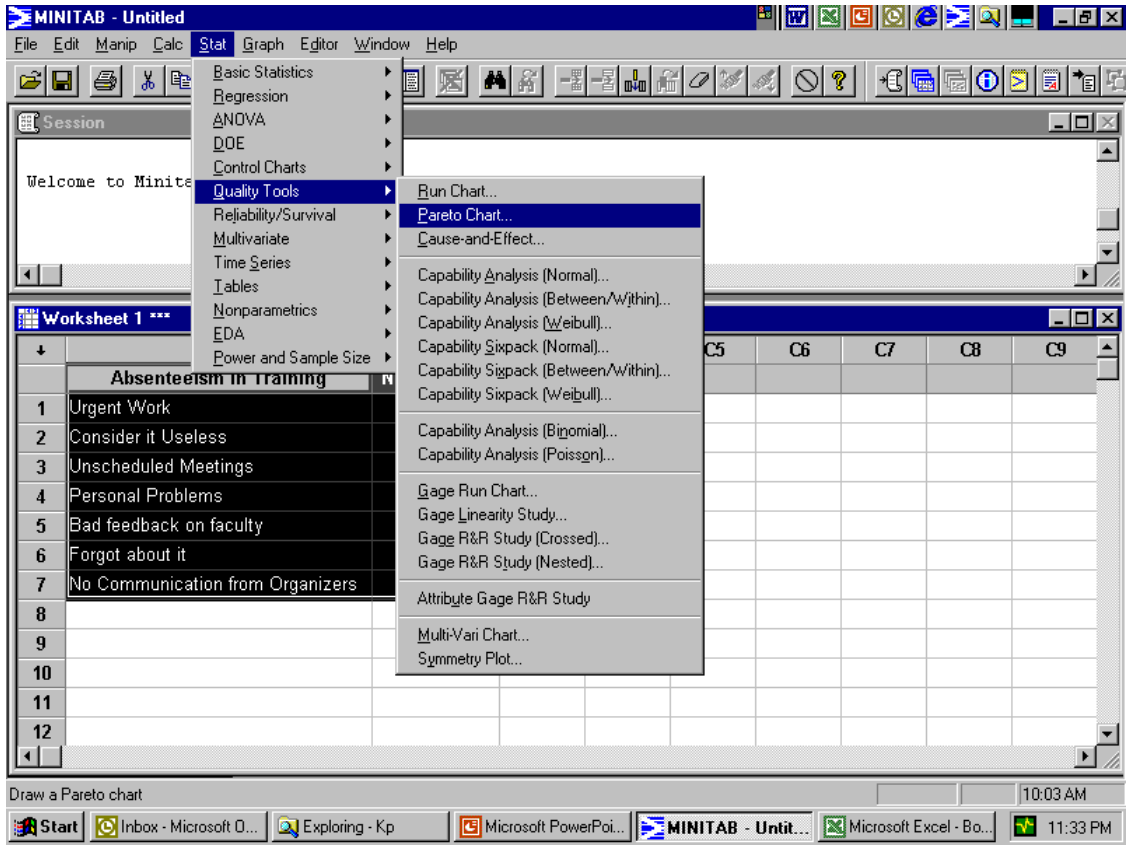
Defect (D)

Anything that results in customer dissatisfaction. Anything that results in a non-conformance. It is an imperfection or deficiency in the output unit with respect to customer specifications

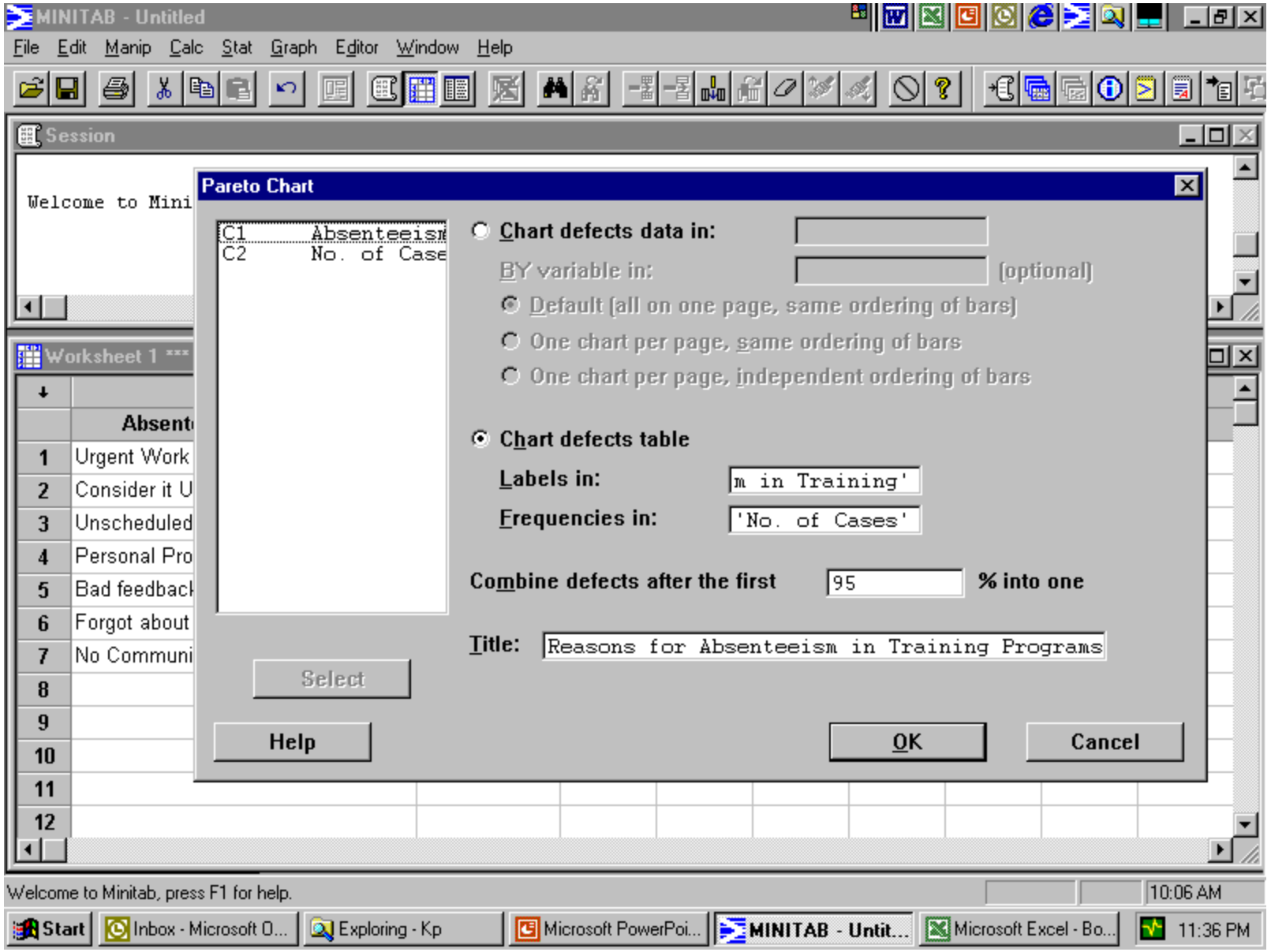
- In a ticket booking example, wrong booking or late delivery is a defect.
- In a pizza delivery case, an order delivered after 40 minutes could be a defect
- In a call center example, phone not picked in 3 rings could be a defect

Green Belt Program

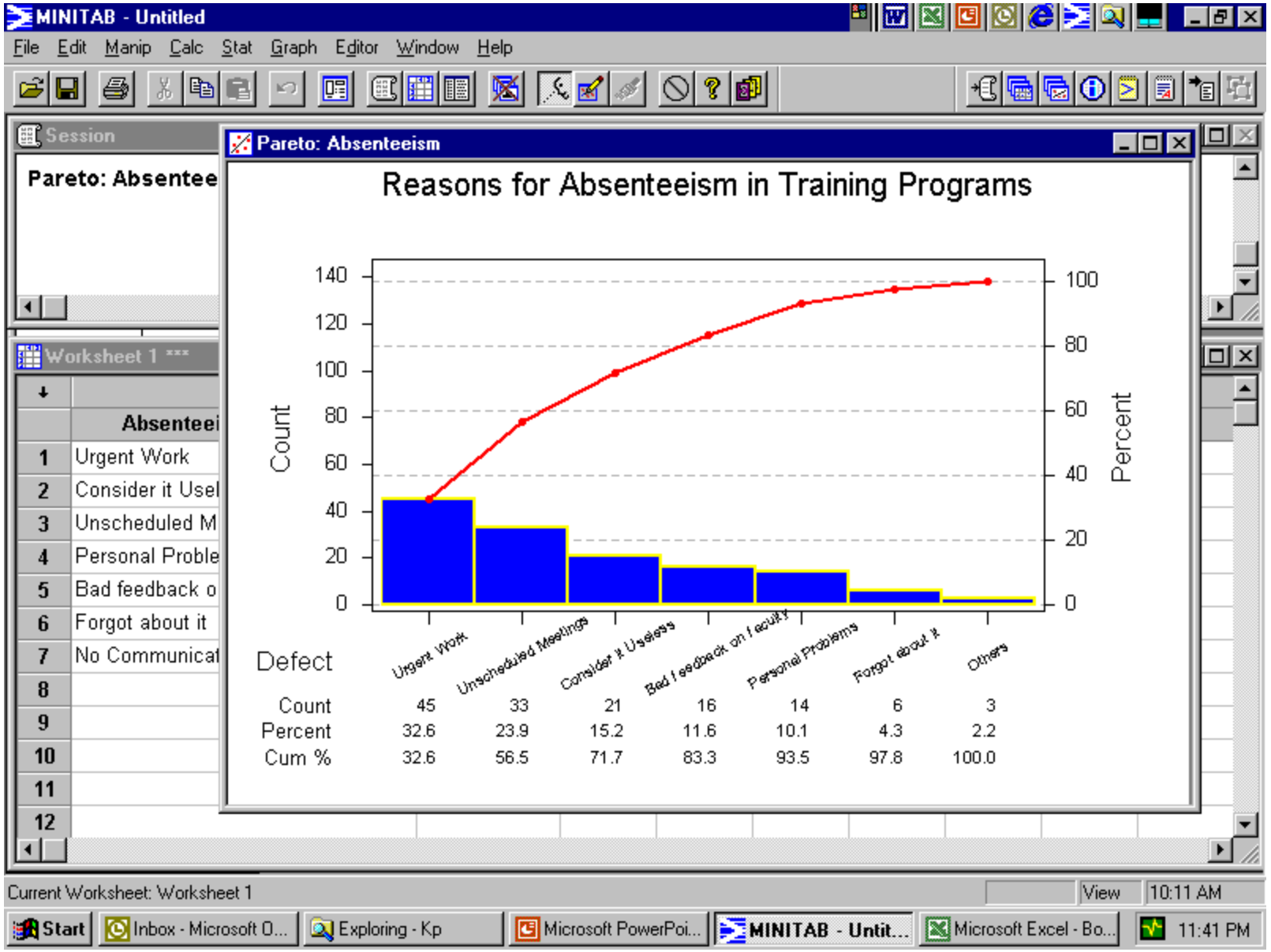
Absenteeism in Training	No. of Cases
Urgent Work	45
Consider it Useless	21
Unscheduled Meetings	33
Personal Problems	14
Bad feedback on faculty	16
Forgot about it	6
No Communication from Organizers	3



Green Belt Program



Green Belt Program



Green Belt Program

$$\text{DPU} = \frac{\text{Number of defects found at any acceptance point}}{\text{Number of units processed through that acceptance point}}$$

Example 1:

Function: Finance

Product: Financial reports

Defect: Entry not accurate

# of Defects: 56

Unit: Each entry

# of Units: 50,000

DPU = 56/50,000 or **.001**

Example 2:

Function: Security

Product: Package inspection at lobby door

Defect: Package not inspected

# of Defects: 380

Unit: Each Package

# of Units: 10,000

DPU = 380/10,000 or **.038**

- When summed, DPU is referred to as *Total Defects Per Unit* or TDU

Green Belt Program

Defect

• The non-fulfillment of intended usage requirements.

• Any non-conformance of a characteristic with specified requirements

• Any condition or characteristic in any supplies or services furnished by the supplier under the contract that is not in compliance with the requirements of the contract.

• e.g. Absence of invoice No. or authorized signature is a defect in an invoice.

Defective

• A unit of product or service having one or more defects

• e.g. An invoice having one or more defects (like absence of Invoice No or Authorized signature) is a defective invoice

Green Belt Program

Opportunity for Error is the number of chances in a process to deviate from the product/process standard.

- **Indicates complexity of process :** Opportunities for defect do not change before and after Project. As the number of opportunities increase, the complexity of the Process increases
- **Need to be independent :** Should one defect trigger-off another defect in the same unit, count only ONCE – ONE OPPORTUNITY per Unit :
- **Measurable :** Opportunities shall be defined from the Customer's perspective. Enables comparison of output quality of dissimilar processes

Concept of Opportunity for Error is applicable only when defect measurement is discrete

Green Belt Program

Unit: each requisition form given to be filled

Defect definition: a form with incomplete details

Requisition Form

Indian Railways

Names:

Date of travel: Class of travel:

From: To:

Train Number: Train Name (Optional):

No of Berths : Berth (Optional): L / M / U / SL / SU

Green Belt Program

Defects per Million Opportunities

Opportunities help to compare dissimilar processes by normalizing the complexity

$$DPO = \frac{\text{Number of defects found at any check-point}}{\left(\text{Number of units processed at that check-point}\right) * \left(\text{Number of opportunities per unit to create a defect}\right)}$$

DPMO = DPO * 1000,000

A multiplier of 1000,000 is used to enable higher sigma measurements

Each DPMO value is unique to a process sigma multiple (Z) – Can be obtained from Z-table or MS Excel

Green Belt Program

Sigma Level vs DPMO					
Sigma	**DPMO**	**Sigma**	**DPMO**	**Sigma**	**DPMO**
0.1	919243.3	2.1	274253.1	4.1	4661.2
0.2	903199.5	2.2	241963.6	4.2	3467.0
0.3	884930.3	2.3	211855.3	4.3	2555.2
0.4	864333.9	2.4	184060.1	4.4	1865.9
0.5	841344.7	2.5	158655.3	4.5	1350.0
0.6	815939.9	2.6	135666.1	4.6	967.7
0.7	788144.7	2.7	115069.7	4.7	687.2
0.8	758036.4	2.8	96800.5	4.8	483.5
0.9	725746.9	2.9	80756.7	4.9	337.0
1	691462.5	3	66807.2	5	232.7
1.1	655421.7	3.1	54799.3	5.1	159.1
1.2	617911.4	3.2	44565.4	5.2	107.8
1.3	579259.7	3.3	35930.3	5.3	72.4
1.4	539827.9	3.4	28716.5	5.4	48.1
1.5	500000.0	3.5	22750.1	5.5	31.7
1.6	460172.1	3.6	17864.4	5.6	20.7
1.7	420740.3	3.7	13903.4	5.7	13.4
1.8	382088.6	3.8	10724.1	5.8	8.5
1.9	344578.3	3.9	8197.5	5.9	5.4
2	308537.5	4	6209.7	6	3.4

DPMO <=> Sigma Convertor

Enter Sigma 3 =>DPMO = **66807**

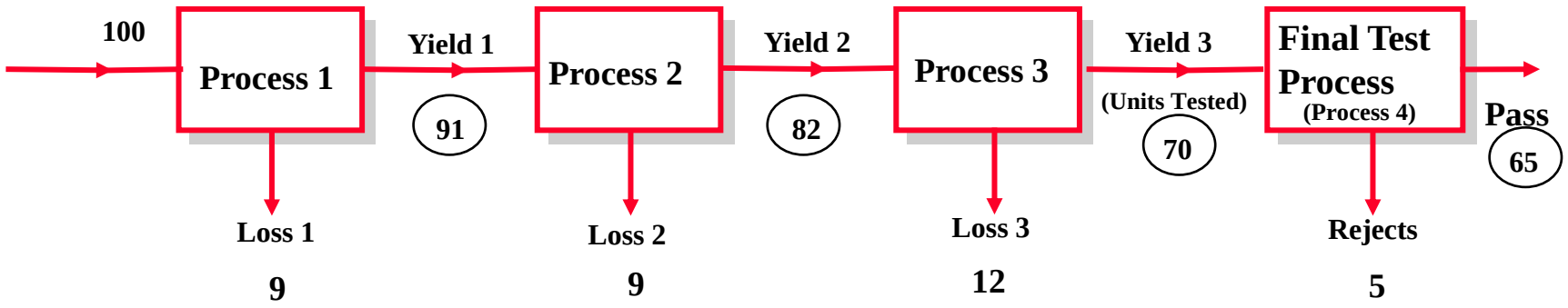
Enter DPMO 62500 =>Sigma= **3.03**

Green Belt Program

A call center company collects the following data on 2000 calls:

Calls received:	2000	Defect:	Not answered with 6 seconds
OFE:	2 (Queue / Not on-line)		
Defective calls:	250		

Calculate the Sigma value (Z) for the overall call efficiency using worksheet.



First Time Yield (Yft)

Yield exclusive of rework

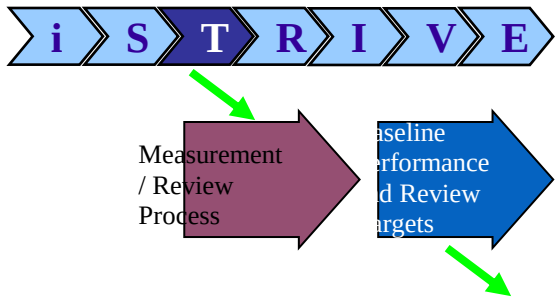
$$= \frac{\text{Units Passed}}{\text{Units Tested}} = \frac{65}{70} = 0.93$$

Rolled Thruput Yield (YTP) $Y_{TP} = P(0) = e^{-DPU}$

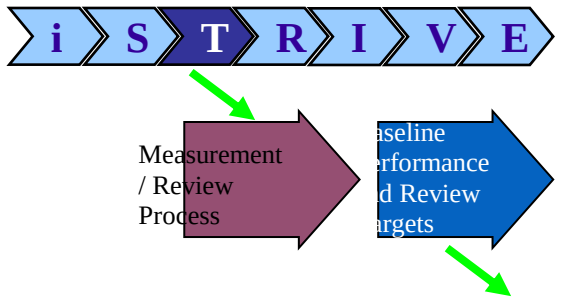
Probability of zero defects = (Yield 1) • (Yield 2) • (Yield 3) . . . .

$$= \left(\frac{91}{100}\right) \cdot \left(\frac{82}{91}\right) \cdot \left(\frac{70}{82}\right) \cdot \left(\frac{65}{70}\right) = 0.65$$

Green Belt Program



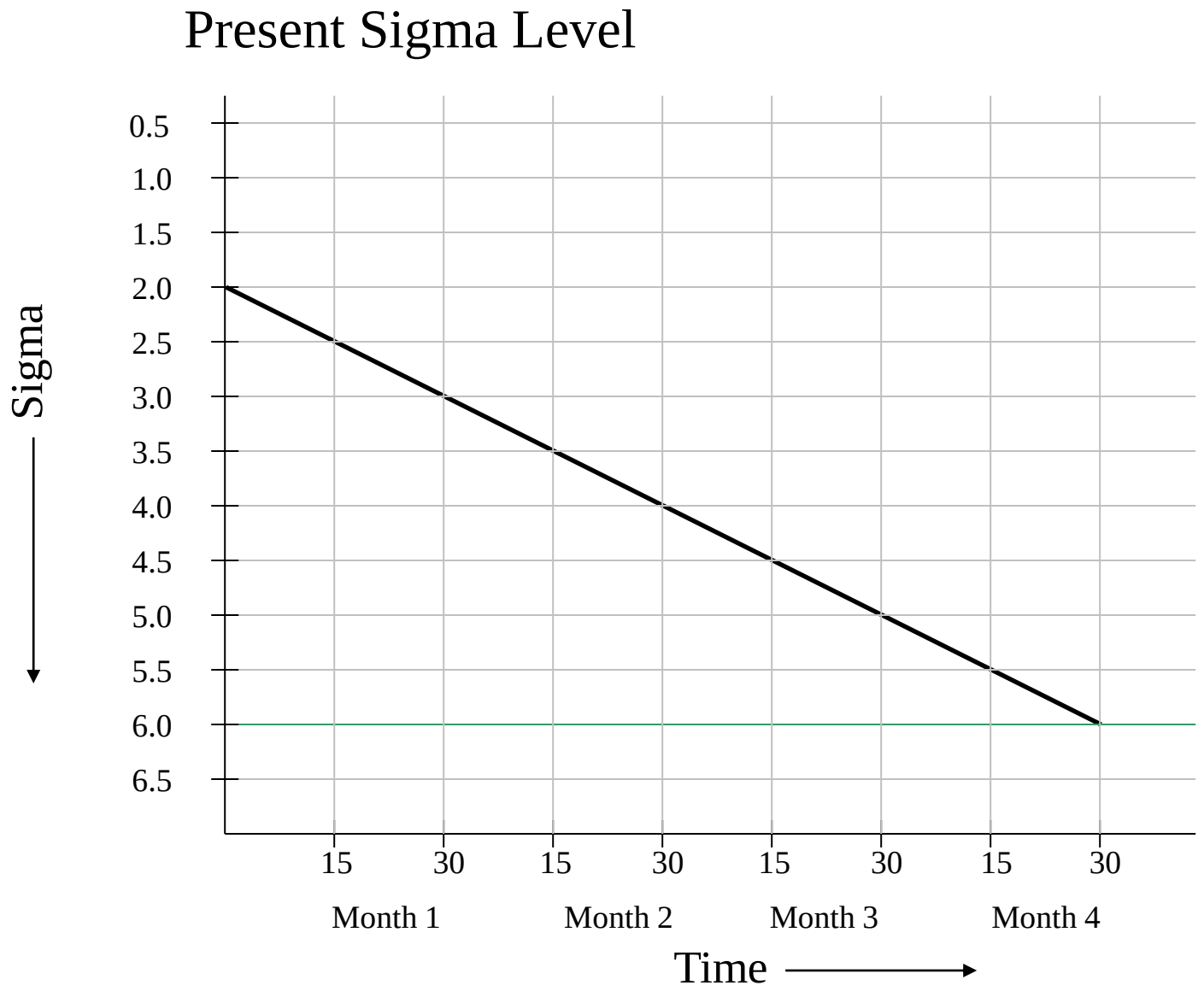
- Generate Data on Project y's
  - Design a simple & efficient data collection check sheet
  - Design a 'suitable' sampling scheme for data collection
  - Review the measurement system R&R
  - Collect the data as per the Plans
- Analyze the Variation for Stability
  - Use simple graphical plots to study trends/process shift & drift
  - Check for Process Stability
    - o Identify the Sporadic Causes in the Process which distort its performance
    - o Eliminate the Sporadic causes and make Process 'Stable'
- Evaluate the Capability of the Process
  - Check the Process for Normality
  - Evaluate the Short term Process Capability

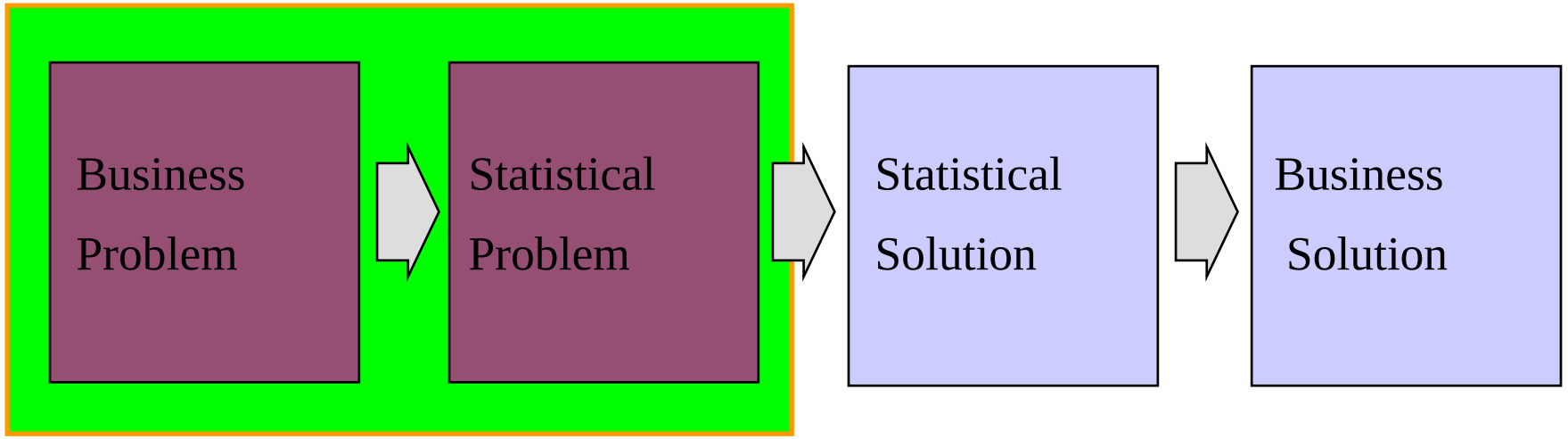


- Review Baseline Performance with set targets
  - Study the Gap between the ‘Set’ Targets and the Current Performance
  - Review the Targets with Stakeholders, if required

- Revise the Targets as per Stakeholders’ needs
- Update Project Charter

Green Belt Program





Data Type	Base Line Values	Target / Goal
Continuous Data	y	y `
	σ	σ`
Discrete Data	DPMO	DPMO`
	DPU	DPU`

Green Belt Program

Questions to be answered

1. How is the PDP measured?

2. How good is the measurement System/Review Process ?

3. Are there any clues about process And measurement system behavior? (Any quick fix solution)

4. How well is the process doing ? (the process capability of 'y')

Deliverables - Phase level

1.To define Statistical Problem

2. Validation of Measurement System

3. Baseline Process performance

Tools

Measurement R&R	**Graphical Plots**	**Control Charts**
Baselining/Z Scorecards	**Histogram**	**Stats Primer**

Green Belt Program

Do's

1. Look at the raw data. What does the Graphical plots tell you?

2. Check for "Trends" in Process Behavior

3. Compare the baseline performance Against PDP Targets

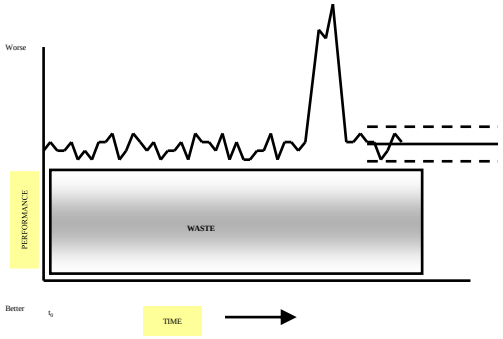
Don'ts

1. Jump on to data transformations. Raw data plots are more informative

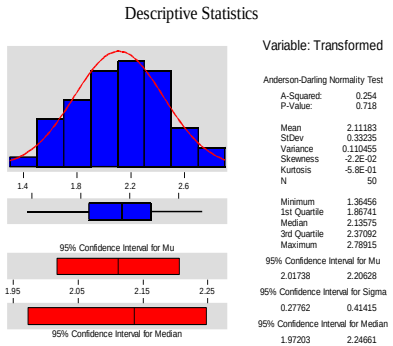
2. Ignore outliers - they are often keys to quick improvements, especially in overall control

Green Belt Program

Measurement System R&R



Track Measurements - PDPs (y's)



Descriptive Statistics

Variable: Transformed

Anderson-Darling Normality Test

A-Squared:	0.254
P-Value:	0.718
Mean	2.11183
StDev	0.33235
Variance	0.110455
Skewness	-2.2E-02
Kurtosis	-5.8E-01
N	50
Minimum	1.36456
1st Quartile	1.86741
Median	2.13575
3rd Quartile	2.37092
Maximum	2.78915

95% Confidence Interval for Mu

2.01738	2.20628

95% Confidence Interval for Sigma

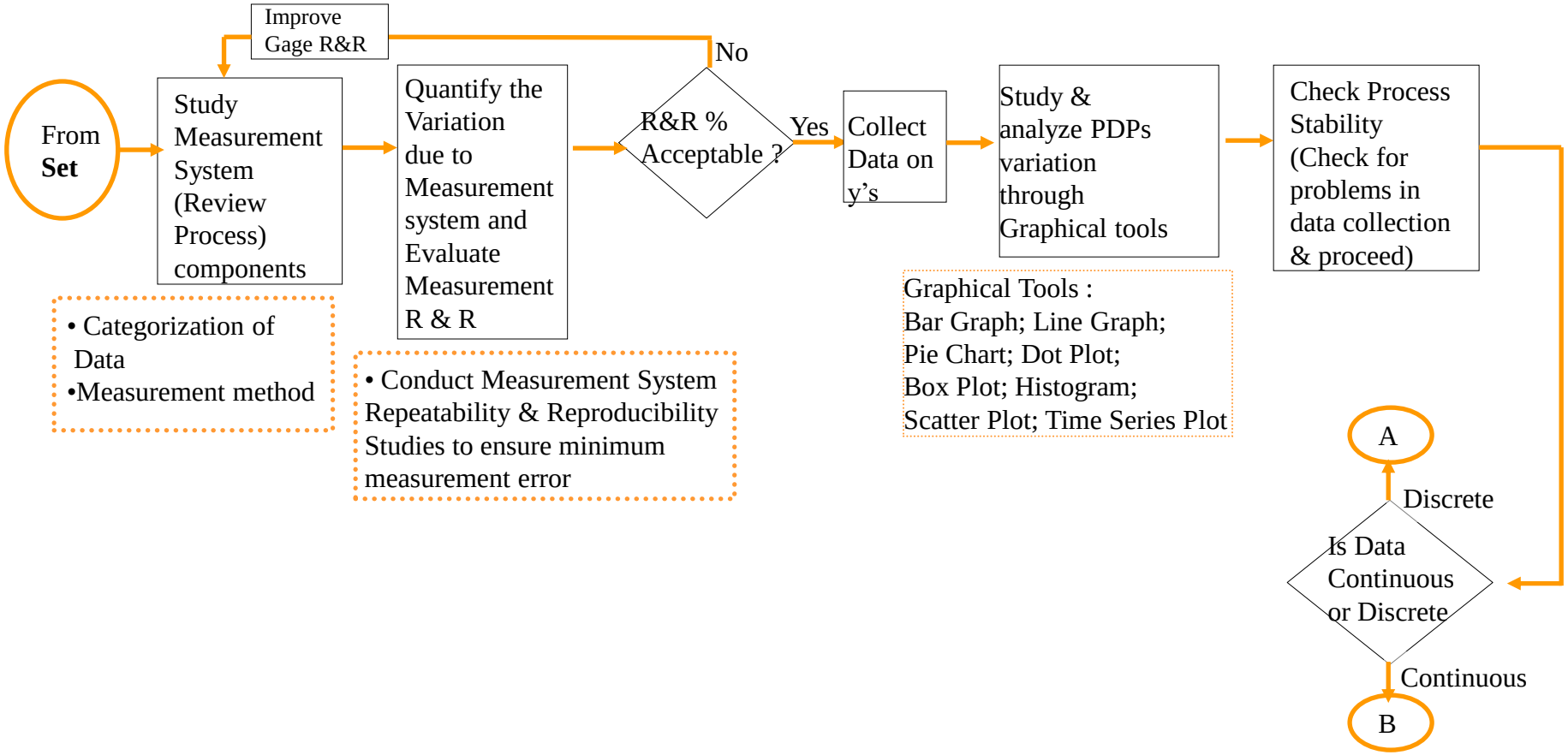
0.27762	0.41415

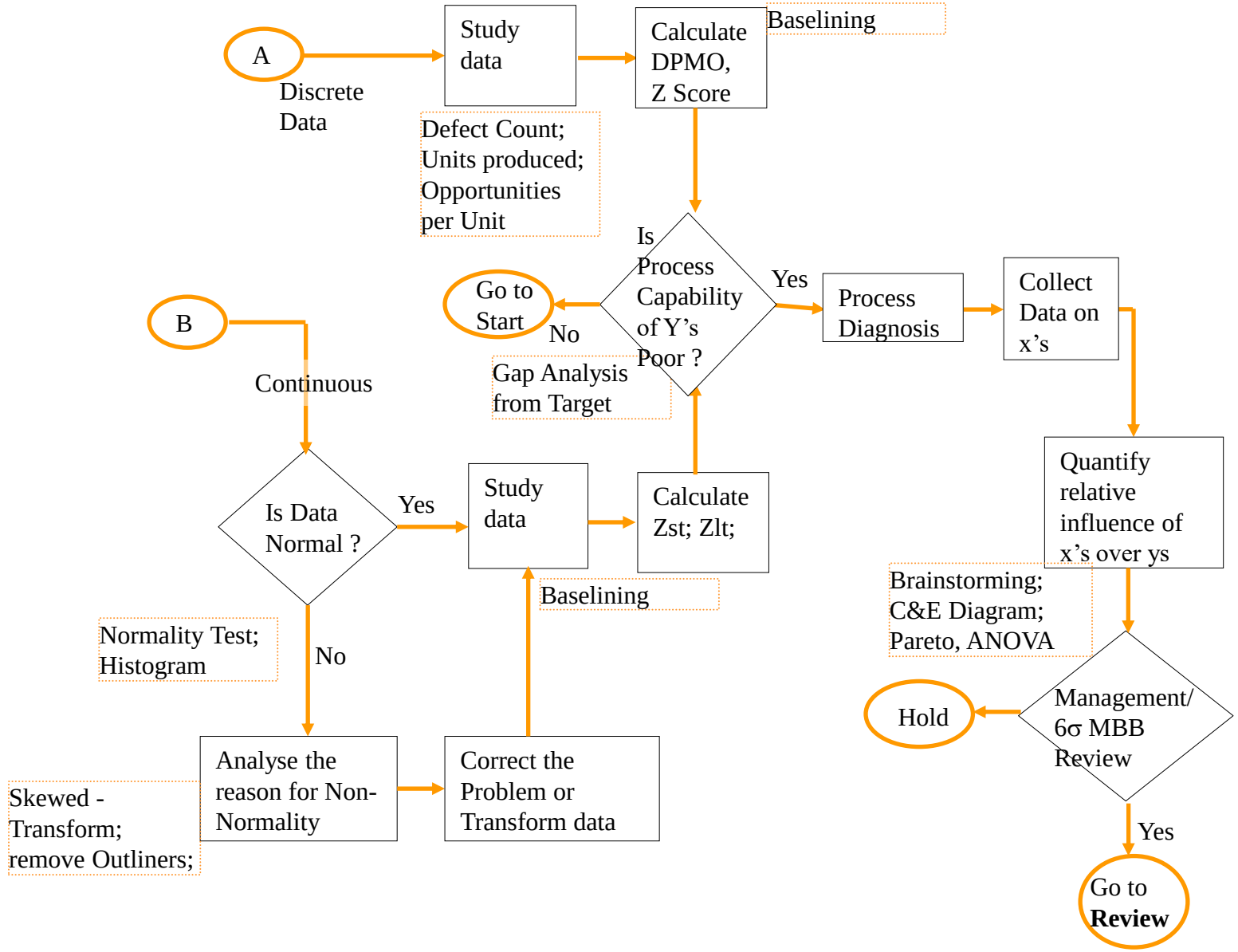
95% Confidence Interval for Median

1.97203	2.24661

Baselining PDP (Project – y) Performance

Track Current Process Capabilities


From **Set**
Study Measurement System (Review Process) components
• Categorization of Data
•Measurement method
Quantify the Variation due to Measurement system and Evaluate Measurement R & R
• Conduct Measurement System Repeatability & Reproducibility Studies to ensure minimum measurement error
R&R % Acceptable ?
No
Improve Gage R&R
Yes
Collect Data on y's
Study & analyze PDPs variation through Graphical tools
Graphical Tools :
Bar Graph; Line Graph;
Pie Chart; Dot Plot;
Box Plot; Histogram;
Scatter Plot; Time Series Plot
Check Process Stability (Check for problems in data collection & proceed)
Is Data Continuous or Discrete
Discrete
A
Continuous
B


A
Discrete Data
Study data
Defect Count; Units produced; Opportunities per Unit
Calculate DPMO, Z Score
Baselining
Is Process Capability of Y's Poor ?
Yes
No
Go to Start
Gap Analysis from Target
Process Diagnosis
Collect Data on x's
Quantify relative influence of x's over ys
Brainstorming; C&E Diagram; Pareto, ANOVA
Management/ 6σ MBB Review
Hold
Yes
Go to Review
B
Continuous
Is Data Normal ?
Yes
No
Normality Test; Histogram
Study data
Calculate Zst; Zlt;
Baselining
Analyse the reason for Non-Normality
Skewed - Transform; remove Outliners;
Correct the Problem or Transform data

Microsoft Word Document

(Invoices)

Microsoft Word Document

(VOR)

Microsoft Word Document

(Delay)

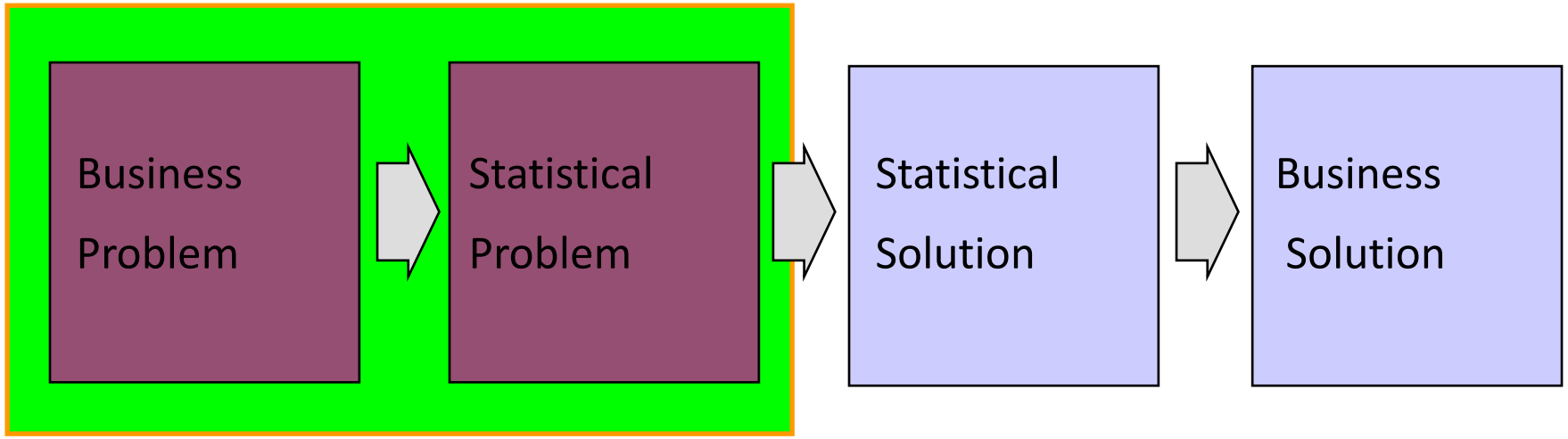
See Appendix for Case Studies

6σ

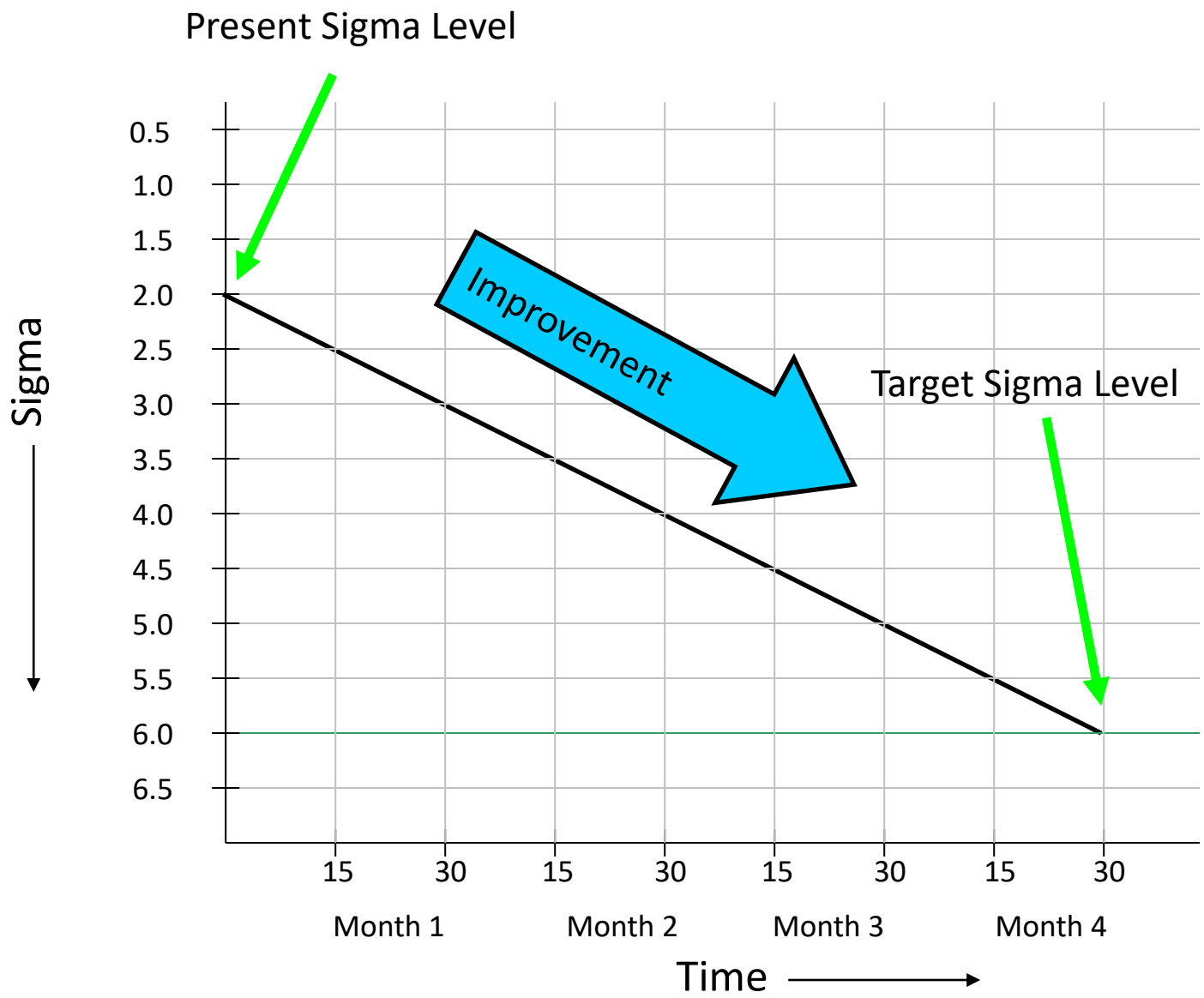
Green Belt Program

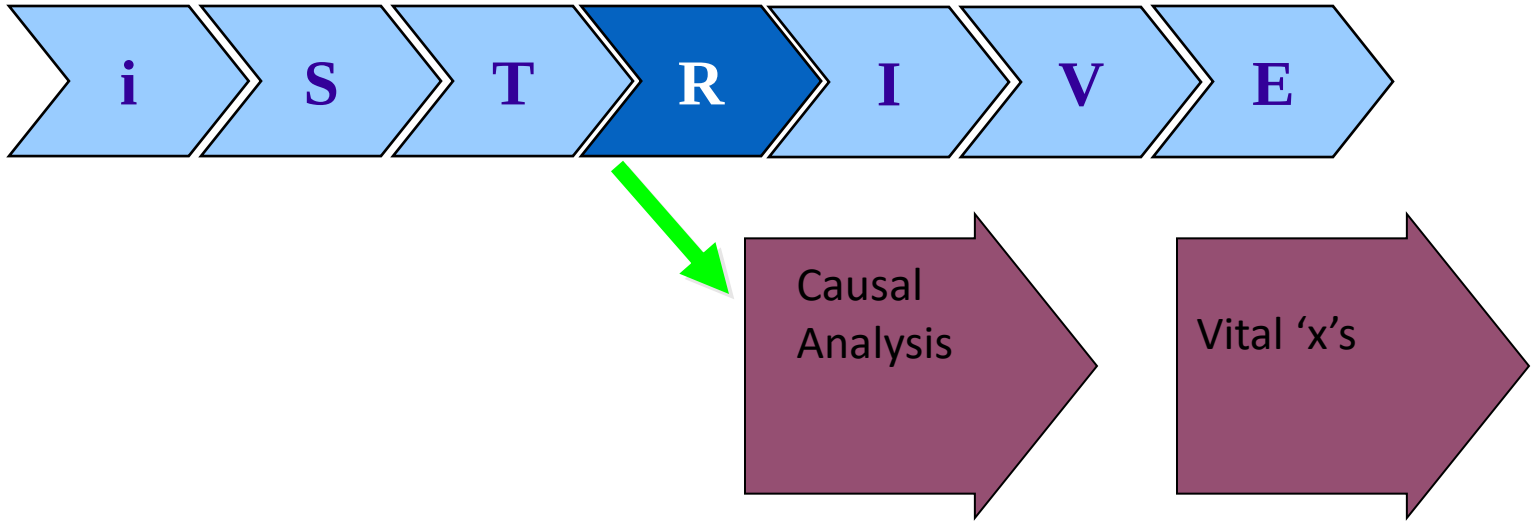
REVIEW

Module 5



Data Type	Base Line Values	Target / Goal
Continuous Data	$\bar{y}$	$\bar{y}$`
	σ	σ`
Discrete Data	DPMO	DPMO`
	DPU	DPU`

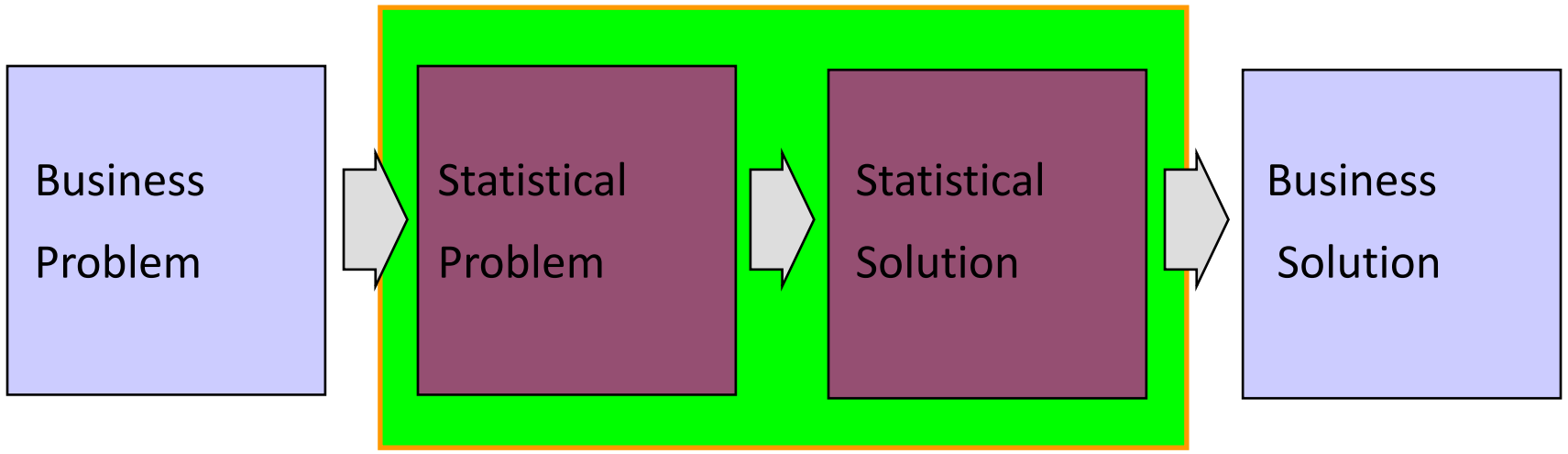

Present Sigma Level
Improvement
Target Sigma Level
Sigma
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
15
30
15
30
15
30
15
30
Month 1
Month 2
Month 3
Month 4
Time



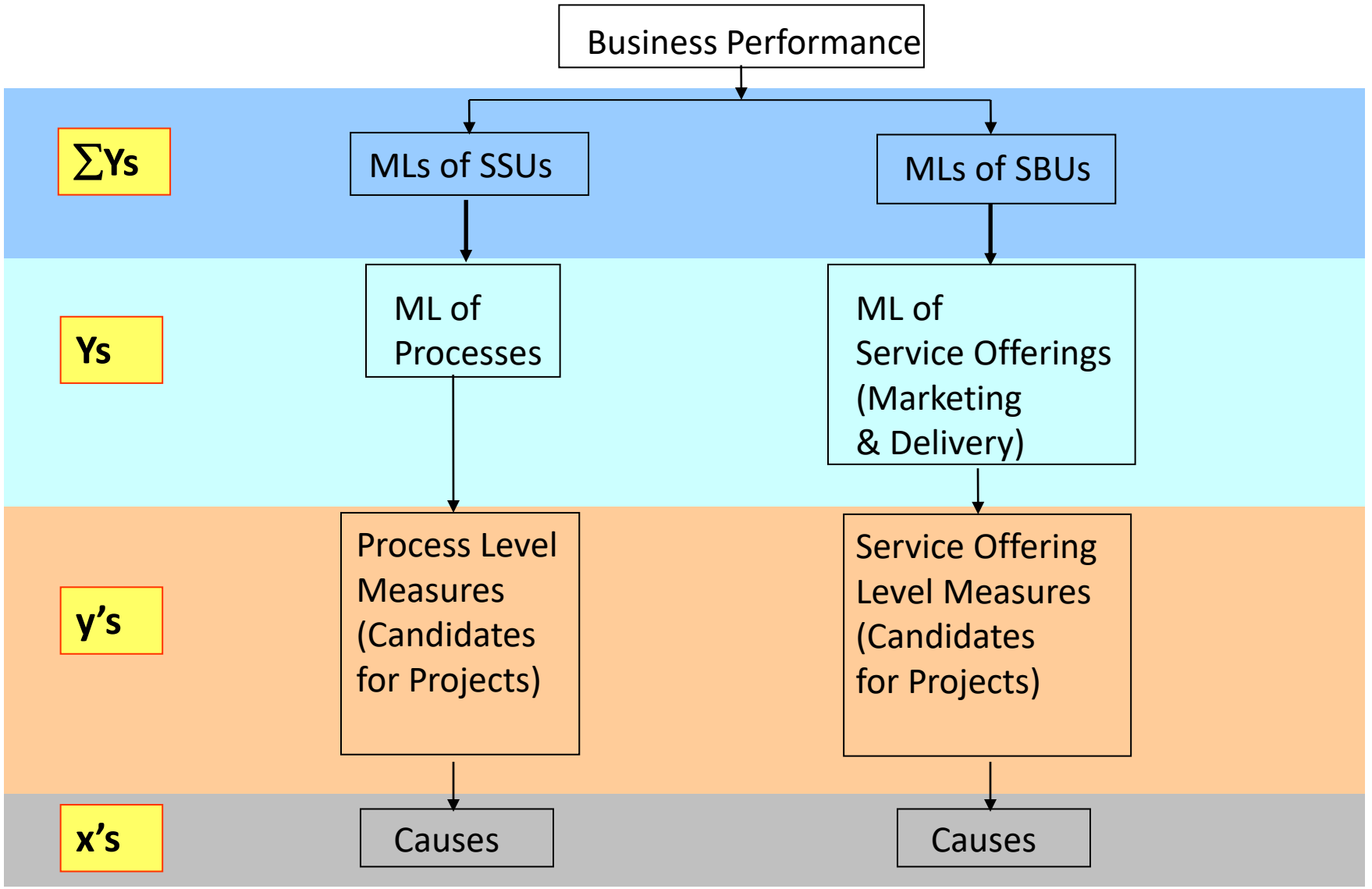
- Develop a detailed Business Process Map
- Analyze the Process Map to identify variation sources
- Determine the x’s (the potential causes)
- Prioritize and collect data on x’s.

- Quantify the relationship y & x’s [y = f(x)]
- Identify the vital x’s from the list of potential x’s
- Validate the vital x’s
- Review Project Risks to deal with vital x’s and develop the Mitigation Plan

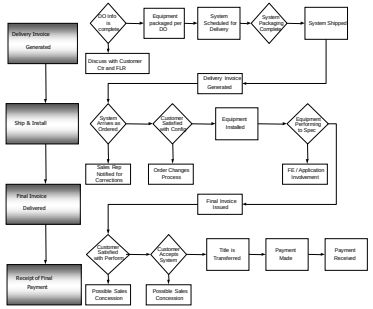
Review the Process to determine Root Causes



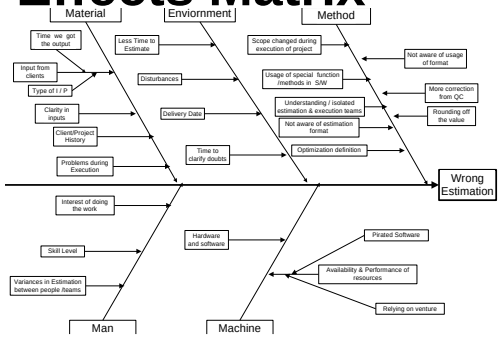
Data Type	Base Line Values	Target / Goal	Vital x's
Continuous Data	$\bar{y}$	$\bar{y}$`	x_i
	σ	σ`	
Discrete Data	DPMO	DPMO`	
	DPU	DPU`	


Business Performance
∑Ys
MLs of SSUs
MLs of SBUs
Ys
ML of Processes
ML of Service Offerings (Marketing & Delivery)
y's
Process Level Measures (Candidates for Projects)
Service Offering Level Measures (Candidates for Projects)
x's
Causes
Causes

Process Map



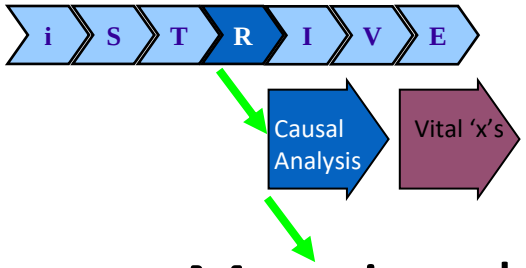
Cause and Effects Matrix



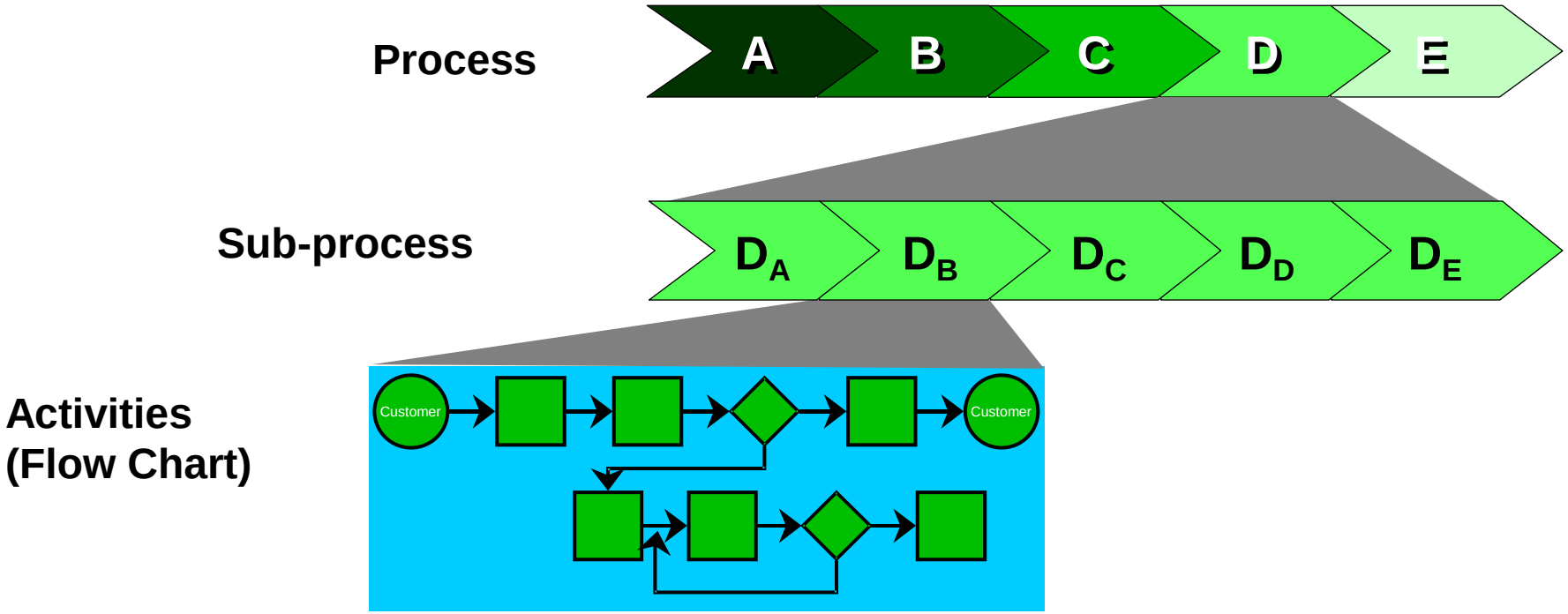
Testing x's

Validated x's

	Process Step	Process Inputs	1 Heavies in Product	2 Lights in Product	3 Moisture in Product	4 Acidity in Product	5	6 Low Capacity From Unit	7 Excessive Downtime	8 Material Losses	9	10 Corrosion of Equipment	11 Poor Reactor Performance	Total
139	Day Tanks	Analysis	10	10	9	9								335
9	Reactor	Cat./HF Ratio						5				8	7	157
7	Reactor	Rxr Temperature						6		5		4	7	149
73	Lights Removal	Condenser Leak			4				8	2		4	1	148
74	Lights Removal	Reboiler Leak			4				8	2		4	1	148
131	Purification	Low Stages	8						8					144
144	Final Storage	Containers	3	2	6	6								140
100	Neutralization	pH Value		6		6						3		138
16	Catalyst Stripper	Pluggage						3	6			5	3	137
111	Drying	Decomposition	2	6		3				2		2		134
39	Drier	Water Carryover			4				6			5	1	132
34	Drier	Molecular Sieve			3	3		2	7	2				125



- Mapping the AS IS Processes to the micro level



Processing step

Decision

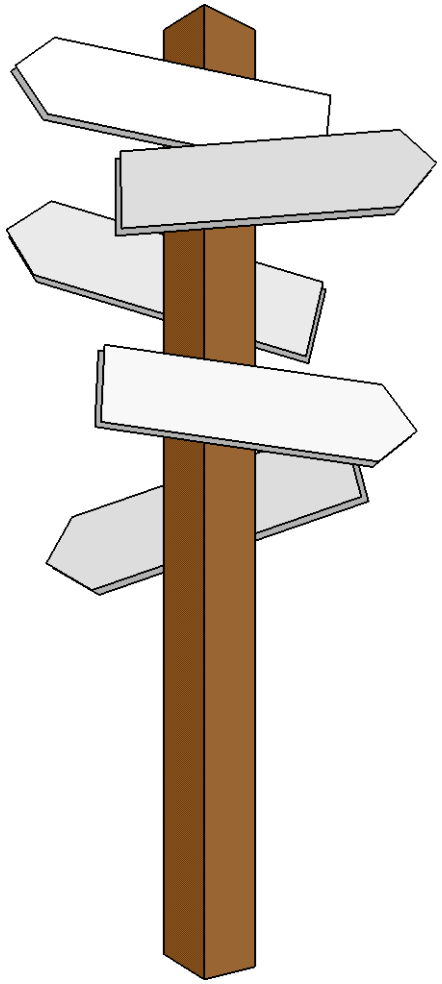
Storage

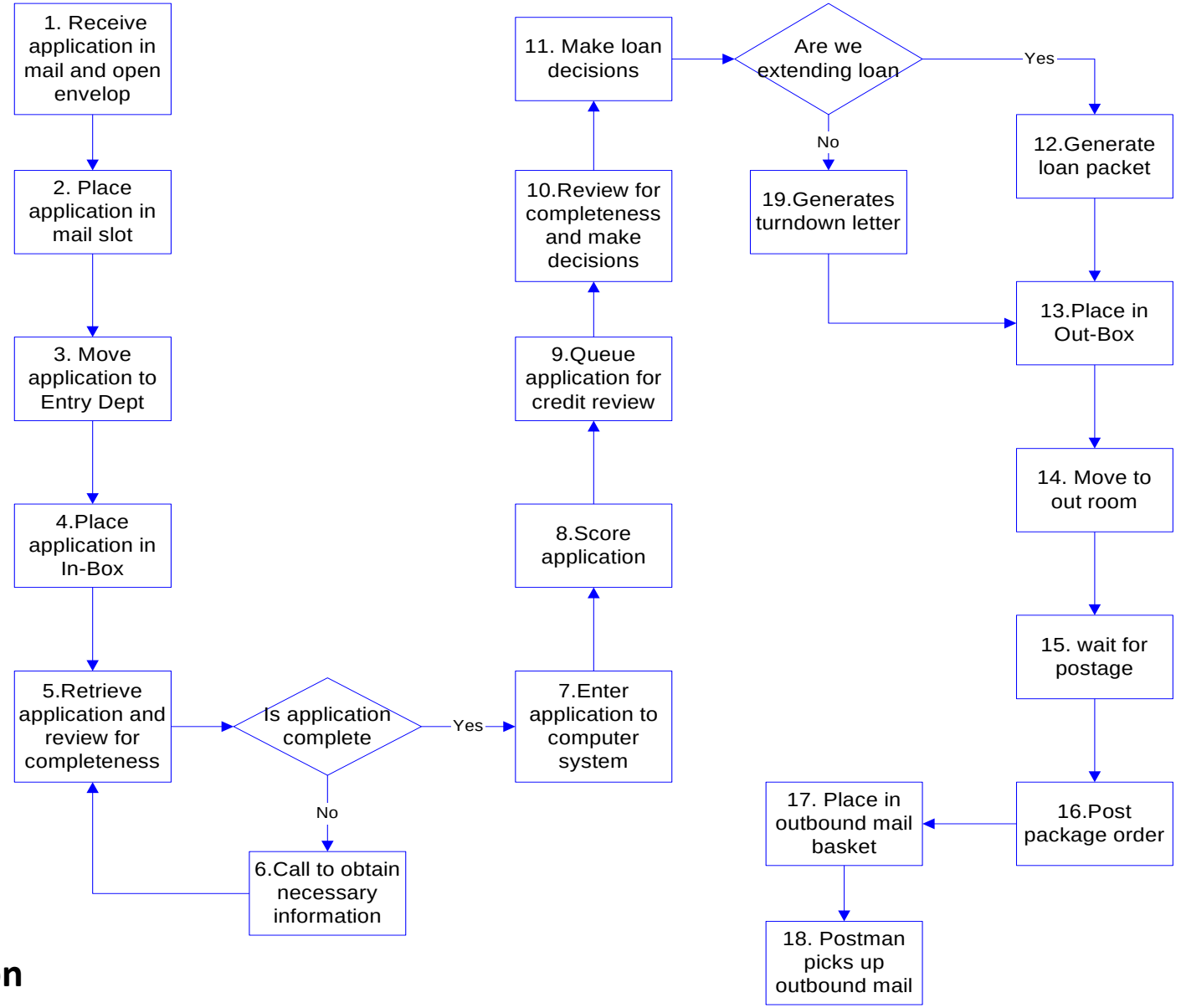
Delay / wait

Predefined process

Start, stop of Process

Process Flow connector





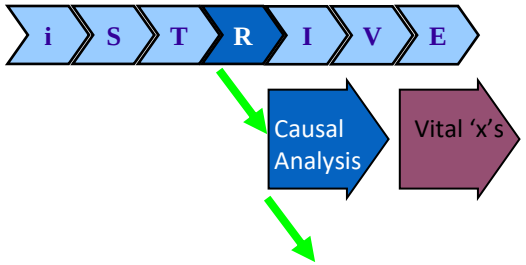
Loan Application

Process Analysis

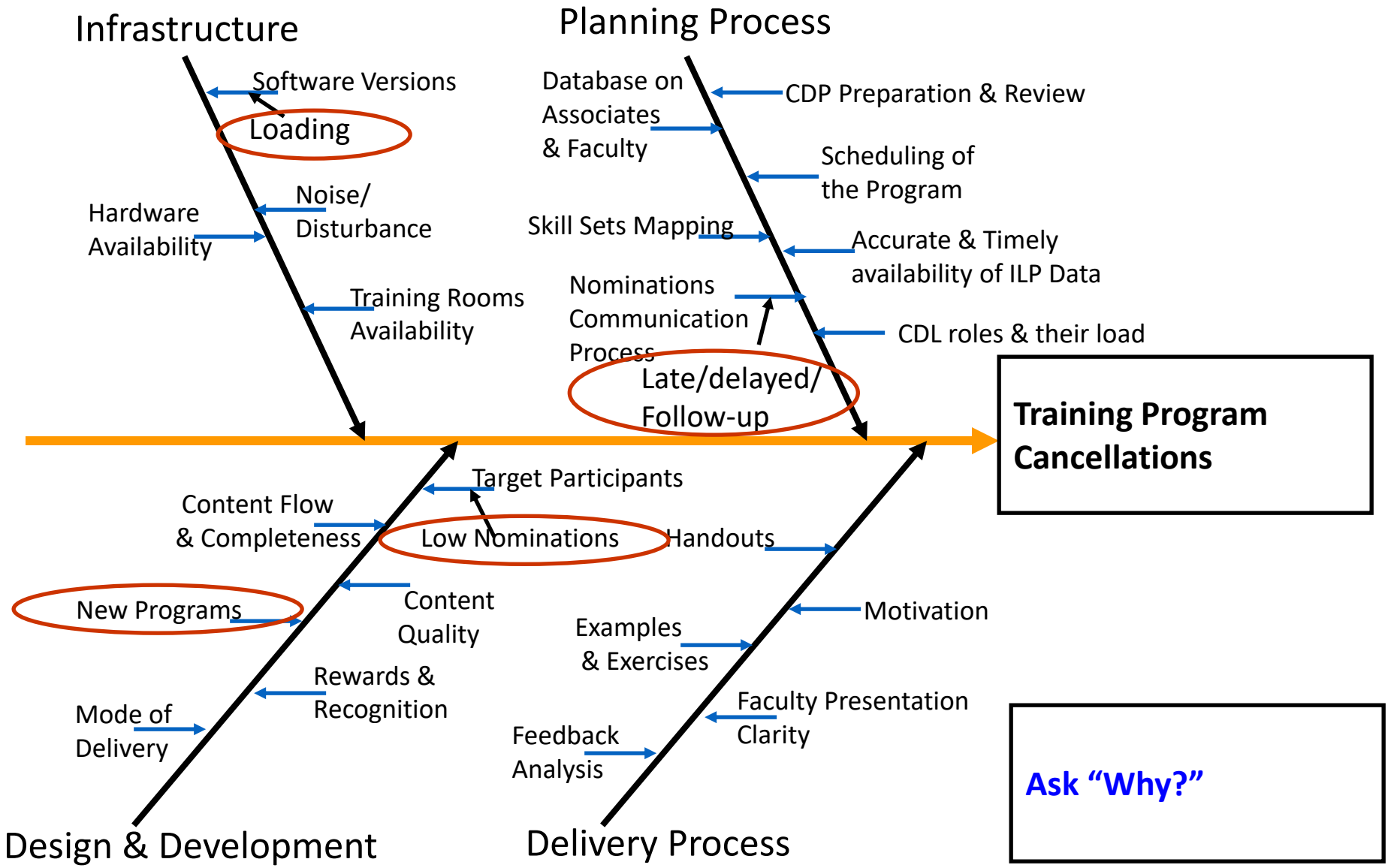
- **Rationalization**
  - Organizational analysis of the process to determine the key processes and assess how these processes interact to meet customer needs
  - It can result in organizational redesign and elimination of some low priority processes

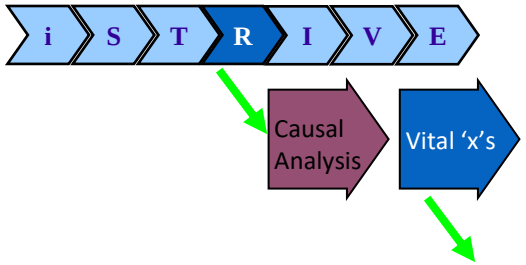
- **Simplification**
  - Systems analytic approach through which teams or individuals study processes to redesign them or eliminate non value added steps

Analyze the Current Process for breakthrough improvements

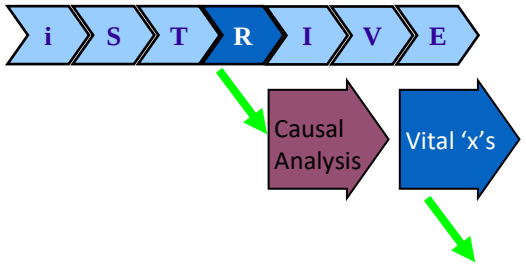


- To identify, explore, and graphically display all of the possible causes related to a problem or condition to discover its root cause(s)
- Enables the team to focus on the content of the problem, not on the history of the problem
- Focuses the team on Causes not symptoms
- Creates a snap shot of the collective knowledge and consensus of a team around a problem


Infrastructure
Software Versions
Loading
Hardware Availability
Noise/ Disturbance
Training Rooms Availability
Planning Process
Database on Associates & Faculty
CDP Preparation & Review
Scheduling of the Program
Skill Sets Mapping
Accurate & Timely availability of ILP Data
Nominations Communication Process
CDL roles & their load
Late/delayed/ Follow-up
Training Program Cancellations
Target Participants
Content Flow & Completeness
Low Nominations
Handouts
New Programs
Content Quality
Rewards & Recognition
Mode of Delivery
Design & Development
Examples & Exercises
Motivation
Faculty Presentation Clarity
Feedback Analysis
Delivery Process
Ask "Why?"



- Identify the vital x's from the list of potential x's, affecting 'y'
- Validate the vital x's - Use Test of Hypothesis to validate x's
- Review Project Risks and develop the Mitigation Plan

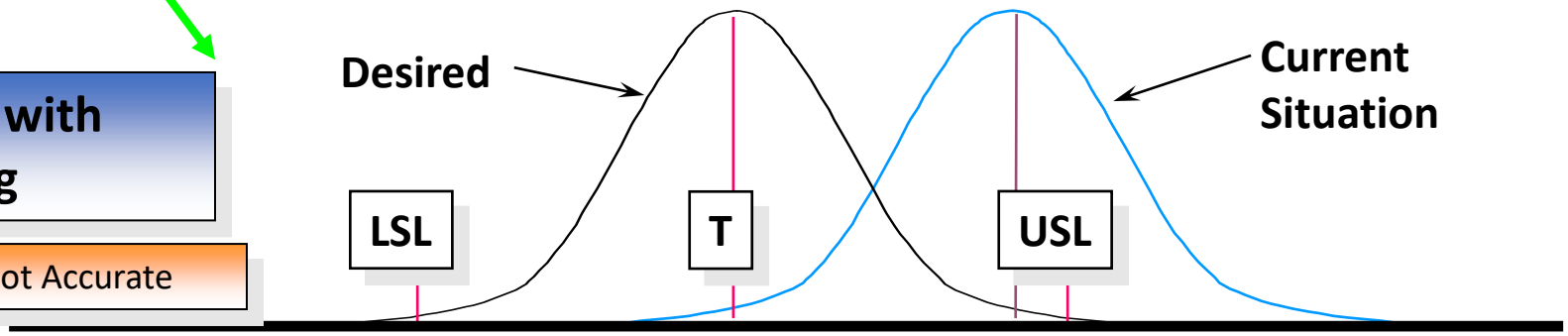


Hypothesis Testing

Central Six Sigma Problems

Problem with Centering

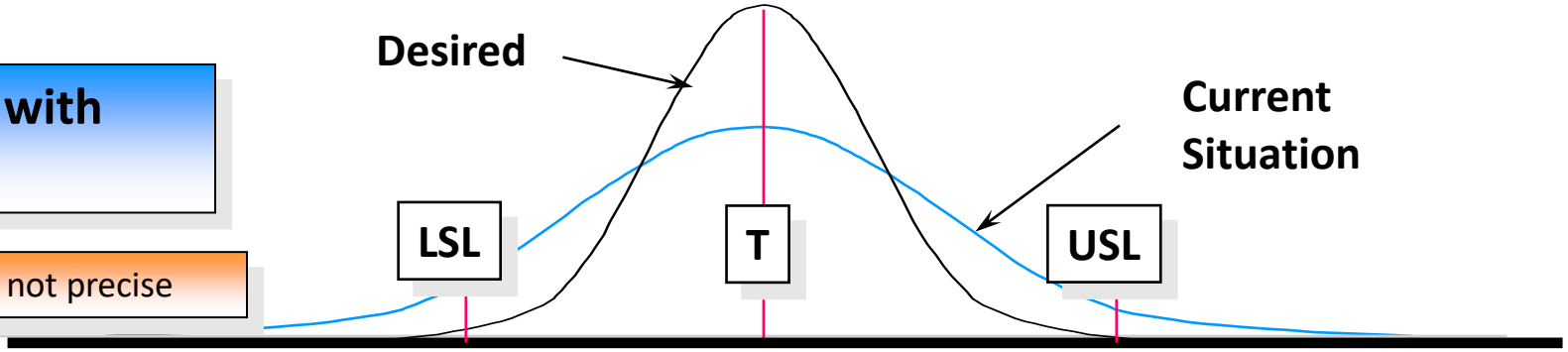
Precise but not Accurate



Are the **Means** of these two processes different?

Problem with Spread

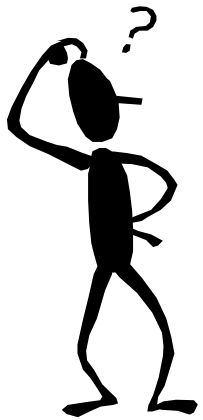
Accurate but not precise



Are the **Variances** of these two processes different?

Hypothesis Testing Enables us to Answer These Questions

Hypothesis Testing



Is there a difference between these two groups?

Is there a difference between these two groups' **means**?

Is there a difference between these two groups' **Variances**?

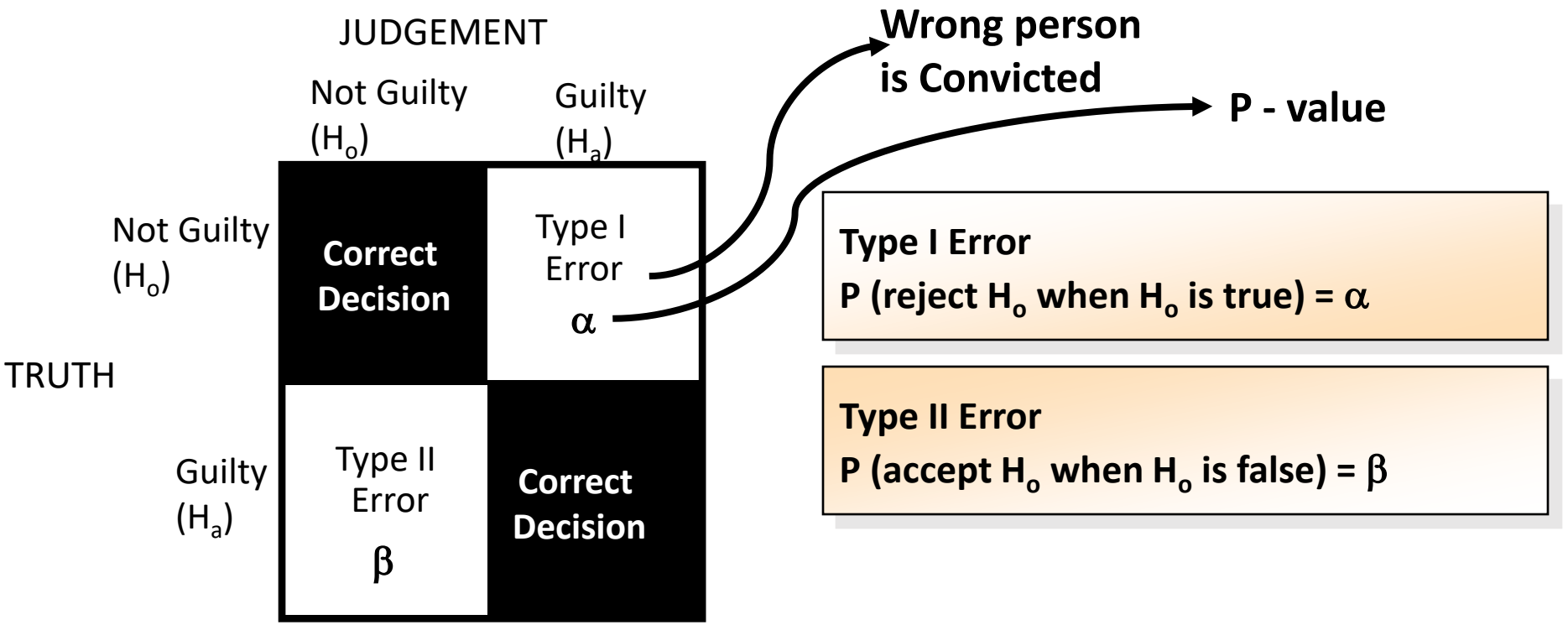
Mechanisms

H_o = Null Hypothesis - There is NO statistical difference between the two groups

H_a = Alternative Hypothesis - There is a statistical difference between the two groups

Hypothesis Testing

Hypothesis Testing Errors



Test "confidence" typically refers to (1-α) = RISK %

Hypothesis Testing

The P-value

- The p-value is the observed probability of making a Type I error.
- Unless there is an exception based on engineering judgment, we usually standardize on a Type I error probability of $\alpha = 0.05$.
- α should be specified before the hypothesis test is conducted

If the p-value > .05 …

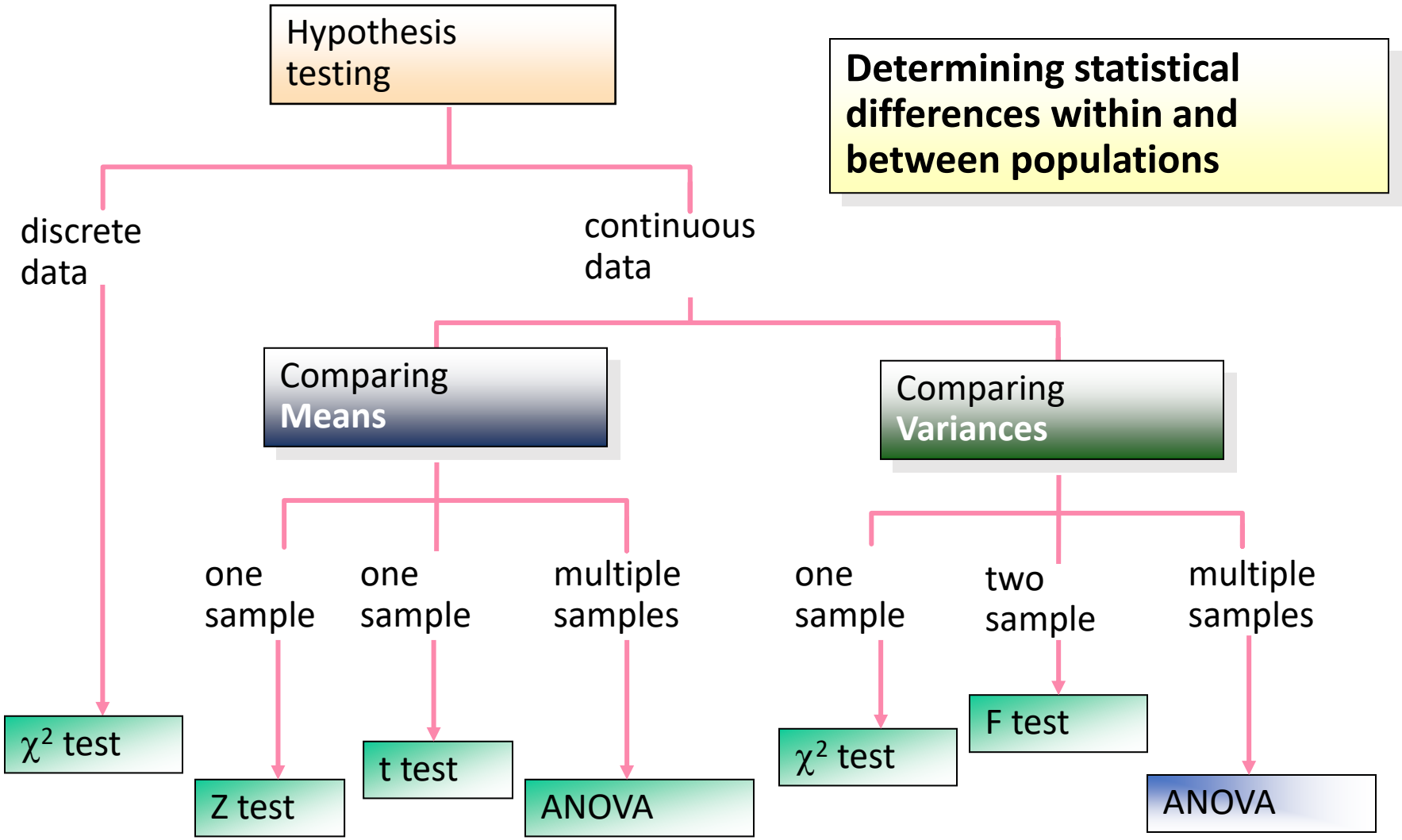
Then H_o is true and there is no difference in the groups

(Accept H_o)

If the p-value < .05…

Then H_o is false and there is a statistically significant difference (Reject H_o)

Hypothesis Testing Road Map



Z-Test

Inference on the Mean of a population, variance known

Example: The Effort Variance is calculated from the projects executed for the last 12 months and the average found to be 20% with standard deviation 3.5%. This is recognized as a critical factor and Management has implemented few process changes improving Estimation & Execution activities with a target to bring-down effort variance. Now fresh data is collected and Effort variance and Test whether the process changes have resulted in improvements ?

Effort variance percentages :

19.8	18.5	17.6	16.7	15.8	10.1	15.4	14.1	13.6
11.9	11.4	7.9	11.4	8.8	7.5	15.4	15.4	19.5
14.9	12.7	11.9	11.4	14.9	15.4	7.5	15.4	11.4
12.7	11.4	17.6	7.9	15.4	11.9	19.5	8.8	11.4
10.1	14.1	13.6	15.8	16.7	19.8	11.9	18.5	

$\bar{x} = 13.71$ $\alpha = 0.05$ $\sigma = 3.5$ $n = 44$

Z-Test

Inference on the Mean of a population, variance known

Solution:

1) The parameter of interest is mean effort variance, m

2) H_0: $\mu = 20$

3) H_1: $\mu < 20$. We want to reject H_0 if Mean Effort Variance is < 20 %

4) $\alpha = 0.05$

5) The test statistic is =>

$$Z_o = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

6) Reject H_0 if $Z_0 < Z_{0.05} = -\mathbf{1.65}$ (Refer Z table)

7)

$$Z_o = \frac{13.71 - 20}{3.5/\sqrt{44}} = -11.921$$

Conclusion: Since $Z_0 = \mathbf{-11.921 < -1.65}$, **Reject H_0** **at $\alpha = 0.05$**

t-Test

Inference on the Mean of a population, variance unknown

If X is normally distributed

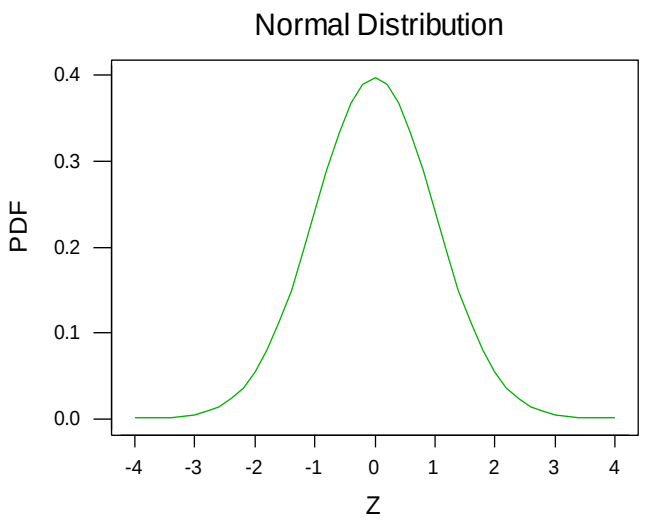
$$Z = \frac{(X - \mu)}{\sigma}$$

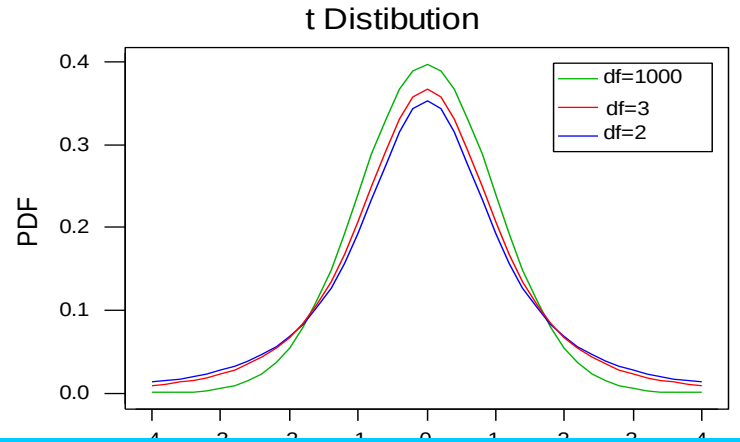
$e.g. \quad Z_{(1-\alpha)} = 1.645 \quad for \quad \alpha = 5\%$

Then $\bar{X}$ is t distributed

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}} \qquad df = n - 1$$

$e.g. \quad t_{(1-\alpha),\, df} = 2.132 \quad for \quad n = 5, \alpha = 5\%$





't' distribution gives p-values for mean comparisons.

t-Test

Confidence Interval on the Mean, Variance unknown

$$\left[\overline{X} - t_{df,\alpha/2} \frac{s}{\sqrt{n}} \right] \leq \mu \leq \left[\overline{X} + t_{df,(1-\alpha/2)} \frac{s}{\sqrt{n}} \right]$$

n = sample size
df = n-1
(1-α) = confidence
s = sample standard deviation

If we take many samples

Then (1-α)% of the time m will be within the CI

t-Test

Inference on the Mean of a population, variance unknown

Example: Fuel Efficiency was calculated from a set of 12 cars of the same model by a manufacturer and the average was found to be 10kmpl. This was considered as too low by the market and hampered sales. The company implemented changes in the manufacturing process and started procuring fuel pumps from another source with a target of improving fuel efficiency to more than 10kmpl . Now fresh data is collected on fuel efficiency. Test whether the process changes have resulted in improvements ?

Fuel Efficiency data:

19.8	18.5	17.6	16.7	15.8	10.1	15.4	14.1	13.6
11.9	11.4	7.9	11.4	8.8	7.5	15.4	15.4	19.5
14.9	12.7	11.9	11.4					

$\bar{x} = 13.71$

$s = 3.55$

Normal Distribution

$\alpha = 0.05$

t-Test

Inference on the Mean of a population, variance unknown

Solution:

1) The parameter of interest is fuel efficiency

2) $H_0: \mu = 10$

3) $H_1: \mu > 10$. We want to reject H_0, if Fuel Efficiency > 10 kmpl

4) $\alpha = 0.05$

5) The test statistic is =>

$$T_0 = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

6) Reject H_0 if $T_0 > T_{0.95,\,21}$ = **1.72** (Refer t-distribution table)

7)

$$T_0 = \frac{13.71 - 10}{3.55/\sqrt{22}} = 4.90$$

Conclusion: Since $T_0 = 4.9 > 1.72$, **Reject H_0 at $\alpha = 0.05$**

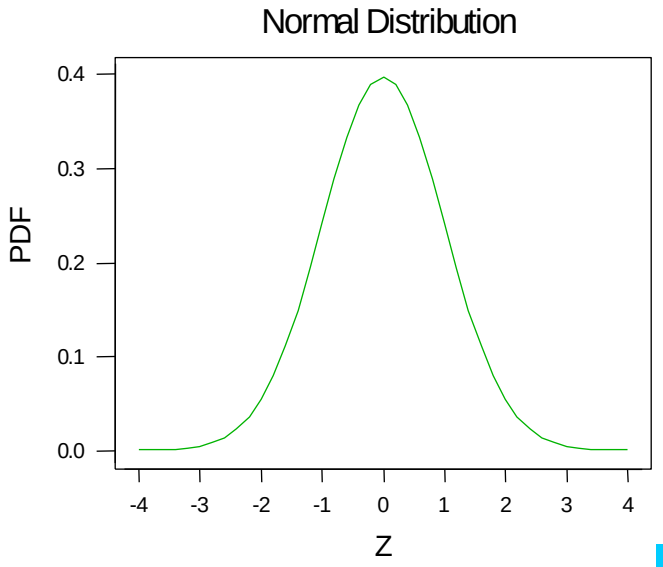
F - Test

"Inference on Variances of two Normal populations"

If X is normally distributed

$$Z = \frac{(X - \mu)}{\sigma}$$

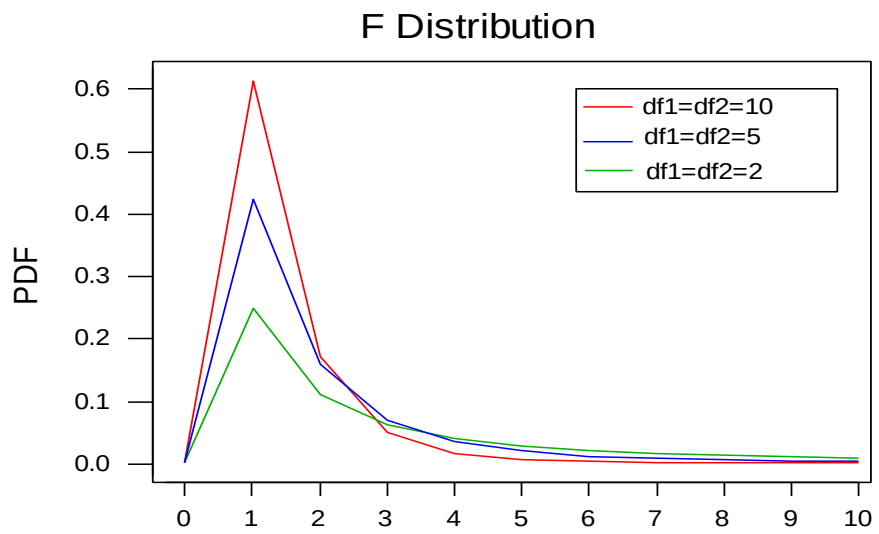
$$e.g. \quad Z_{(1-\alpha)} = 1.645 \quad for \quad \alpha = 5\%$$



Then s_1^2 / s_2^2 is F distributed

$$F = \frac{s_1^2}{s_2^2} \qquad df_1 = n_1 - 1 \quad df_2 = n_2 - 1$$

$$e.g. \quad F_{(1-\alpha),df_1,df_2} = 6.388 \quad for \quad n_1 = n_2 = 5, \alpha = 5\%$$



F distribution allows

two-sample variance comparisons.

Two-sided Two Sample F-Test

Test

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_a : \sigma_1^2 \neq \sigma_2^2$$

Compute

$$F = \frac{s_1^2}{s_2^2} \qquad df_1 = n_1 - 1 \qquad df_2 = n_2 - 1$$

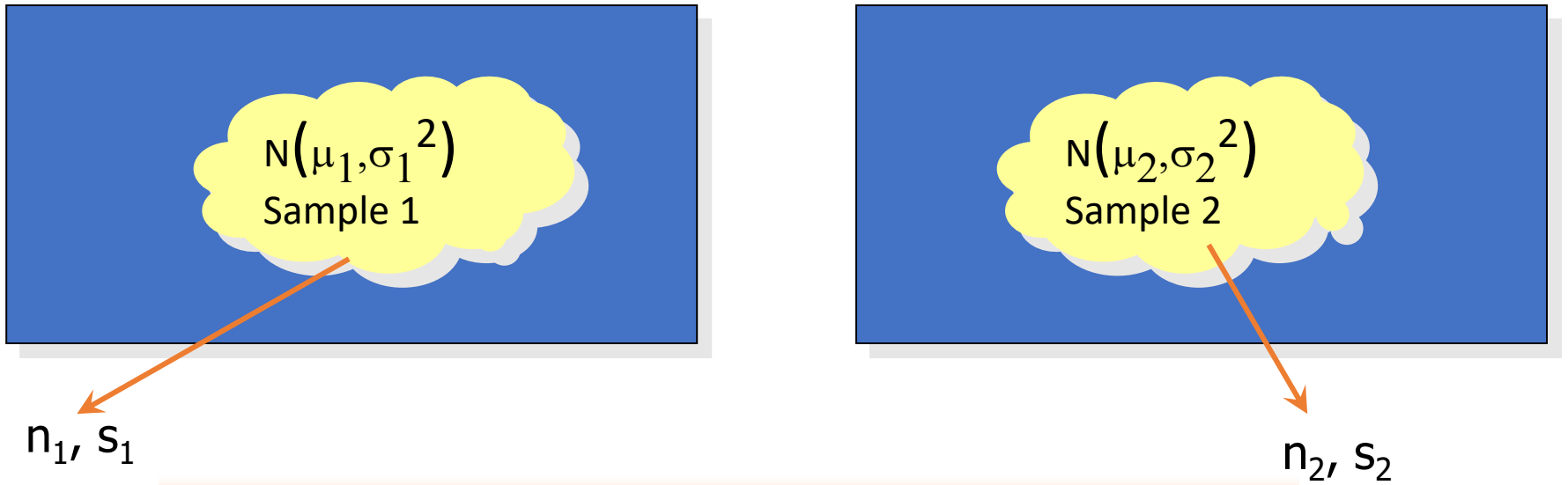
for $s_1 > s_2$

$$F_{critical} = F_{(1-\alpha),\, df1,\, df2}$$

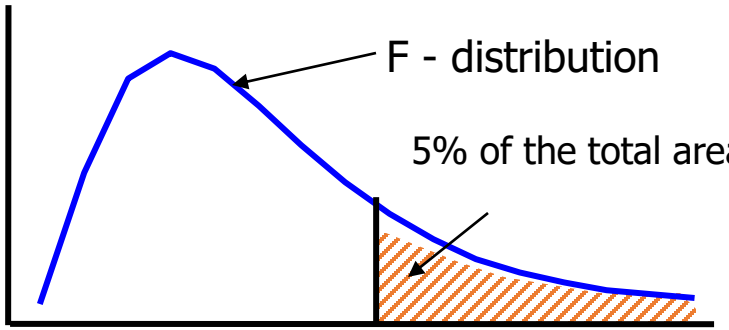
If $F_{computed} > F_{critical}$

Then reject H_0

F - test: Comparing Variances(contd.)



Calculate $F_{calc} = s_1^2 / s_2^2$ such that F>1 and compare to F_{crit}



If $F_{calc} > F_{crit}$, variances s_1^2 and s_2^2 are different !

Example: F- test

Inference on Variances of two Normal populations

Energy consumption is recorded across different locations of Satyam. The variation in the Energy consumption is a critical factor. Low variability is desirable for ensuring smooth operations & control expenditure Two different locations are studied to find out which one is superior with minimum variation in energy consumption. 20 weeks data is collected from these two locations.

Sample std. Deviations of the energy consumption, location-wise is :

s_1 = 196 kWh; s_2 = 213 kWh

Is there any evidence to indicate that either location is better?

Use α = 0.05

Example: F- test

Solution:

1) The parameter of interest is variances of energy consumption s_1^2 and s_2^2
   Assume that energy consumption is a normal random variable for both locations
2) $H_0: s_1^2 = s_2^2$ $H_a: s_1^2 \neq s_2^2$
3) $a = 0.05$
4) The test statistic is $=> f_0 = s_1^2 / s_2^2$
5) Since $n_1 = n_2 = 20$, we will reject $H_0: s_1^2 = s_2^2$ if $f_0 > f_{0.025,19,19} = 2.53$
6) Computations:
   Since $s_1^2 = (1.96)^2 = 3.84$ and $s_2^2 = (2.13)^2 = 4.54$, the test statistic is $f_0 = s_2^2/s_1^2 =$ **1.176**

Conclusion: Since $f_0 =$ **1.176** $< f_{0.025,19,19} =$ **2.53**, we cannot reject the null hypothesis. There is no strong evidence that either location results in less variance of energy consumption

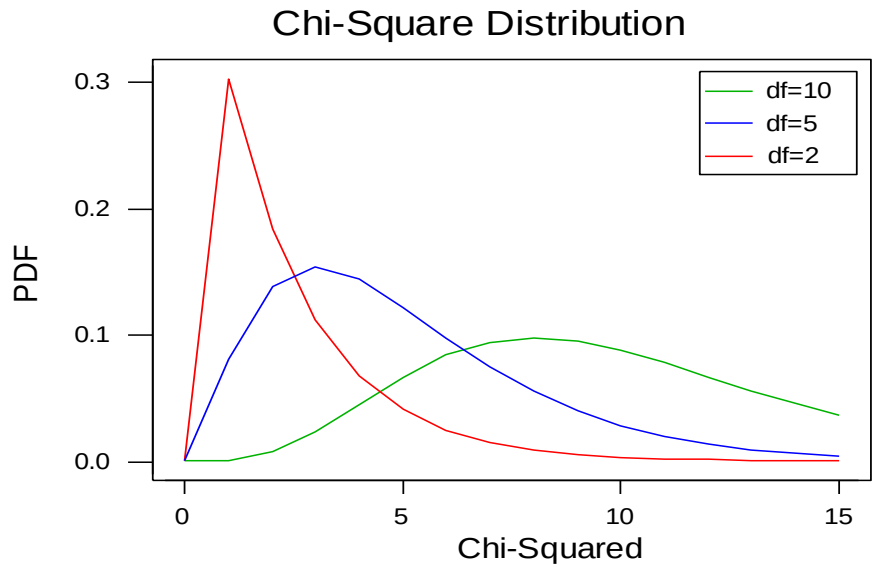
Chi-Squared Distribution

Inference on Proportions (Discrete Data)

Hypothesis tests on the Proportions from different populations

$H_o : p_1 = p_2 = p_3 = p_4 = p$

$H_a : p_1 \neq p_2 \neq p_3 \neq p_4 \neq p$



The test statistic is given by

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$ where

f_o = observed frequency

f_e=expected frequency

Example

- Suppose, we are analyzing the centuries scored by an Indian Player in India and abroad, against 4 countries. We develop a contingency table as below:

Number of centuries scored in India & abroad, against various countries					
	England	Australia	Pakistan	Sri Lanka	
In India	3	6	5	7	21
Abroad	2	2	4	2	10
	5	8	9	9	31
	Total				

- Let’s hypothesize that proportion of centuries in India or abroad is equal.

H_0 : “Proportion of centuries against any country is same”

H_a : “Proportion of centuries scored against different countries are different”

It identifies whether there are any significant differences
But
Does not identify where the differences are ?

- If null hypothesis is true, then there should be a relationship / dependence among the various centuries scored
- We can calculate the *expected frequencies assuming there is a relationship*
- Expected frequency for each of the observed frequency

- For example Observed frequency of 3 against England in India, would convert to expected frequency as $\left(\frac{\text{Row total}}{\text{Grand total}}\right) * \left(\frac{\text{Column total}}{\text{Grand total}}\right) * \text{Grand total}$ $(21 / 31) * 5 = 3.39$

		Number of centuries scored in India & abroad, against various countries				
		England	Australia	Pakistan	Sri Lanka	
In India	Observed	3	6	5	7	21
	Expected	3.39	5.42	6.10	6.10	21
Abroad	Observed	2	2	4	2	10
	Expected	1.61	2.58	2.90	2.90	10

Sample Statistics

Estimated
Population Parameters

- Combining all the information

f_o	f_e	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
3	3.39	0.15	0.04
6	5.42	0.34	0.06
5	6.10	1.20	0.20
7	6.10	0.82	0.13
2	1.61	0.15	0.09
2	2.58	0.34	0.13
4	2.90	1.20	0.41
2	2.90	0.82	0.28
			1.36

$$\chi^2_{\text{Calculated}} = \sum \frac{(f_o - f_e)^2}{f_e} = 1.36$$

- For chi-square distribution, degrees of freedom are calculated as per the number of rows & columns in the contingency table

Degrees of Freedom	=	(Number of rows – 1)	*	(Number of columns – 1)

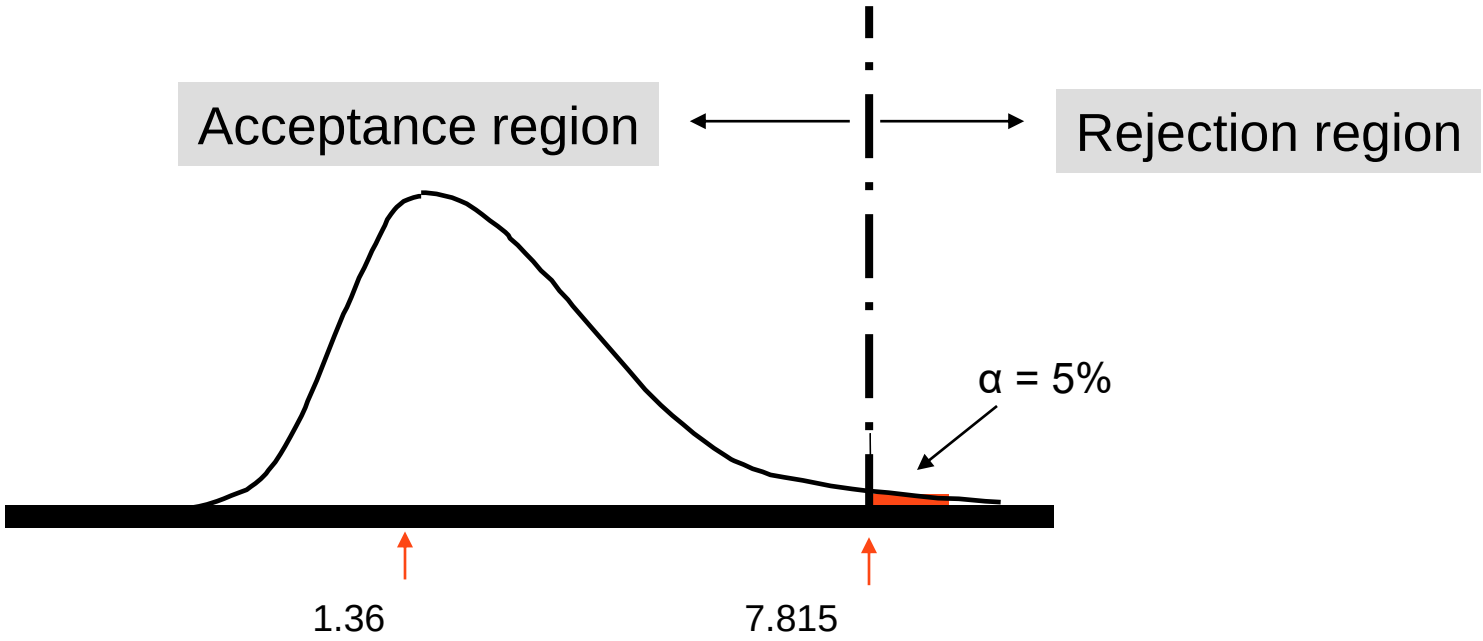
- For this example, degrees of freedom:

  (2 - 1) * (4 – 1) = 3

- Assuming $\alpha = 5\%$, we can look up the Chi-square distribution in Tables (in the end) & arrive at χ^2 Critical = 7.815

- For our example, χ^2 Critical = 7.815

  χ^2 Calculated = 1.36



If $\chi^2_{Calculated} > \chi^2_{Critical}$ then reject H_0

- When calculated value is greater than critical value, it means there is more than $(1 - \alpha)$ % chance that there is a significant difference between given sample proportions & estimated population proportions

- That means samples given do not belong to the same population, & hence we reject the null hypothesis

- When calculated value is less than critical value, it means that samples given belong to the same population, hence we accept the null hypothesis

Chi-Square Test

Expected counts are printed below observed counts

	ENG	AUS	PAK	SRI	Total
1	3	6	5	7	21
	3.39	5.42	6.10	6.10	
2	2	2	4	2	10
	1.61	2.58	2.90	2.90	
Total	5	8	9	9	31

Chi-Sq = 0.044 + 0.062 + 0.197 + 0.134 + 0.093 + 0.131 +
0.414 + 0.281 = 1.356

DF = 3, **P-Value = 0.716**

- Since p-value is more than α, we accept the null hypothesis

Questions to be answered

1. What could cause the symptoms ?

2. What critical input variables or causes (x's) affect the process output (y's) ?

3. How much impact will each 'x' cause on y ?

4. What is the statistical significance of the x's?

5. Have we collected the appropriate data ?

Deliverables - Phase level

1.Detailed Process Map

2.Identification of vital x's

3. Validated x's

Tools

Brainstorming	Cause & Effect Diagram	Process Mapping
Test of Hypothesis	Correlation & Regression	FMEA
ANOVA	Scatter Plots	

Do's

1. Begin with potential causes - Use structured brainstorming, cause-effect diagrams, FMEA and other tools to develop theories of potential causes. These become your theories of test - begin with the most likely

2. Use the information derived from the statistical analysis of the process measures, - to guide discussions for selection of most likely causes

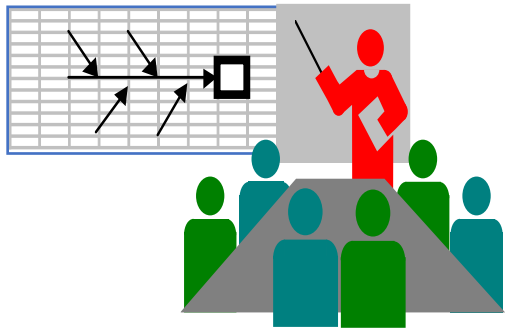
3. Review the Problem for further categori-sation into sub-problems, if any.

Don'ts

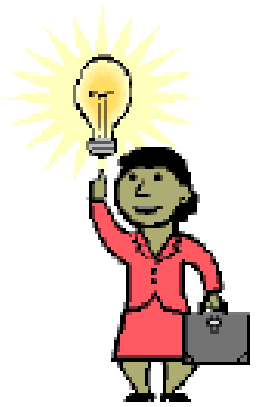
1. Ignore the rigor of Micro level process mapping.

2. Assume you know the cause because of earlier information - there is a potential for spurious correlation's of events and relationships must be examined

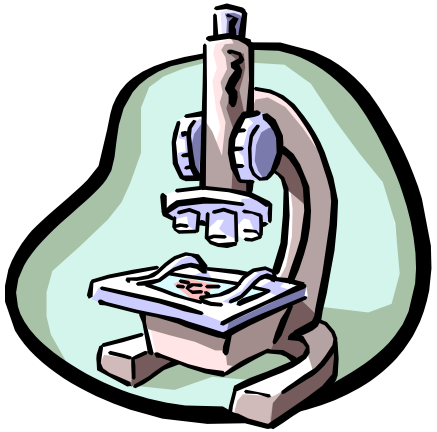
3. Assume the "causes" based on numerical analysis of the data and the summary statistics only - Look for graphical plots also to confirm the hypothesis.



Causal Analysis

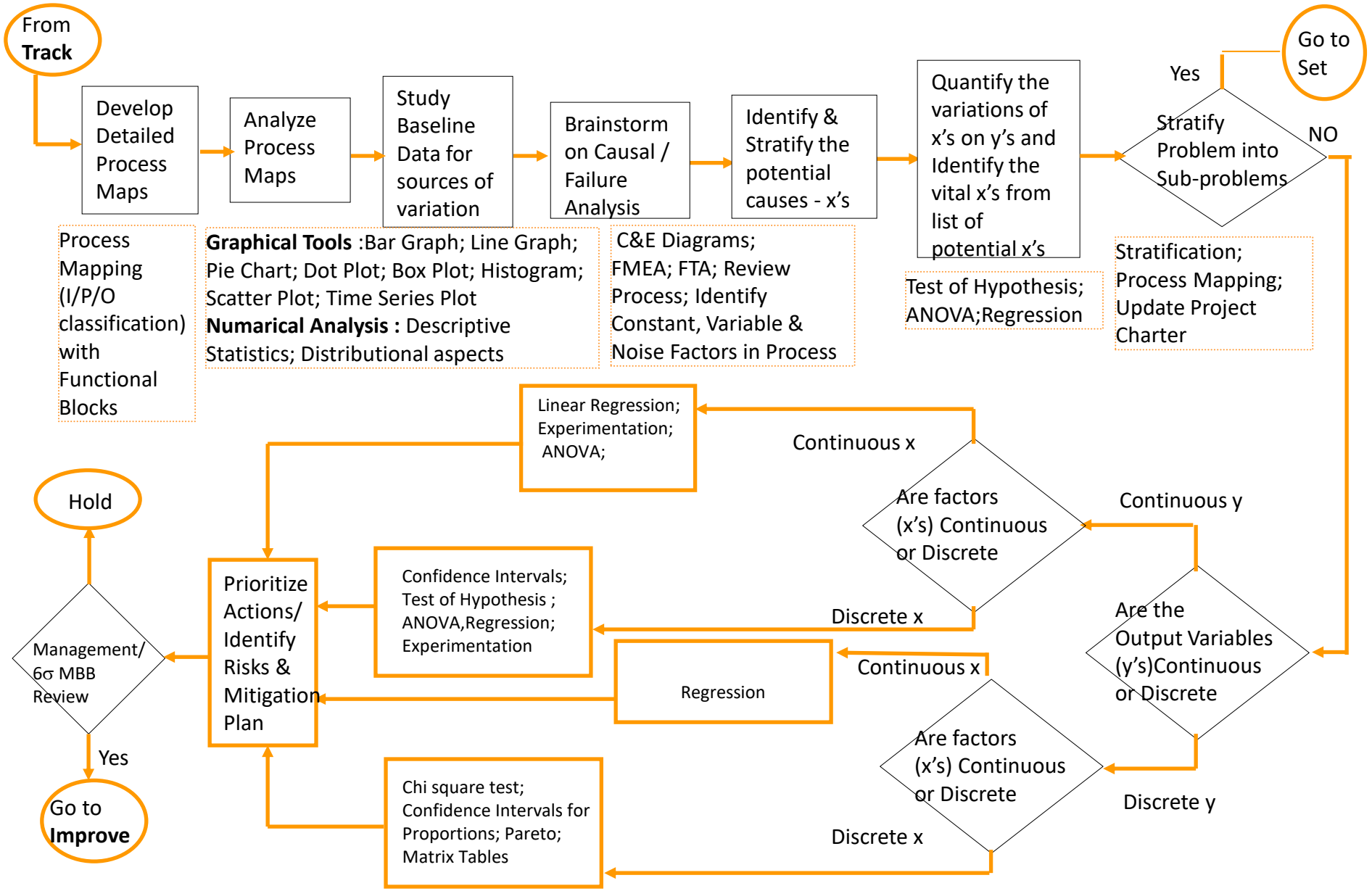


Identify
Potential x's



Identify and Validate
Vital x's

Review the Process to determine Root Causes


From **Track**
Develop Detailed Process Maps
Analyze Process Maps
Study Baseline Data for sources of variation
Brainstorm on Causal / Failure Analysis
Identify & Stratify the potential causes - x's
Quantify the variations of x's on y's and Identify the vital x's from list of potential x's
Stratify Problem into Sub-problems
Yes
NO
Go to Set
Process Mapping (I/P/O classification) with Functional Blocks
Graphical Tools :Bar Graph; Line Graph; Pie Chart; Dot Plot; Box Plot; Histogram; Scatter Plot; Time Series Plot
Numarical Analysis : Descriptive Statistics; Distributional aspects
C&E Diagrams; FMEA; FTA; Review Process; Identify Constant, Variable & Noise Factors in Process
Test of Hypothesis; ANOVA;Regression
Stratification; Process Mapping; Update Project Charter
Linear Regression; Experimentation; ANOVA;
Continuous x
Are factors (x's) Continuous or Discrete
Continuous y
Are the Output Variables (y's)Continuous or Discrete
Discrete x
Continuous x
Regression
Are factors (x's) Continuous or Discrete
Discrete y
Discrete x
Chi square test; Confidence Intervals for Proportions; Pareto; Matrix Tables
Confidence Intervals; Test of Hypothesis ; ANOVA,Regression; Experimentation
Prioritize Actions/ Identify Risks & Mitigation Plan
Management/ 6σ MBB Review
Hold
Yes
Go to **Improve**

Microsoft Word
Document

(Analyzing the Root Cause of Problems)

See Appendix for Case Study

Microsoft Word Document

(Has the Spare Parts Delivery Improved)

Microsoft Word Document

(Spare Parts Delivery Revisited)

Microsoft Word Document

(Investigating Variability in Spare Parts Delivery)

Microsoft Word Document

(Is the Western Region the best)

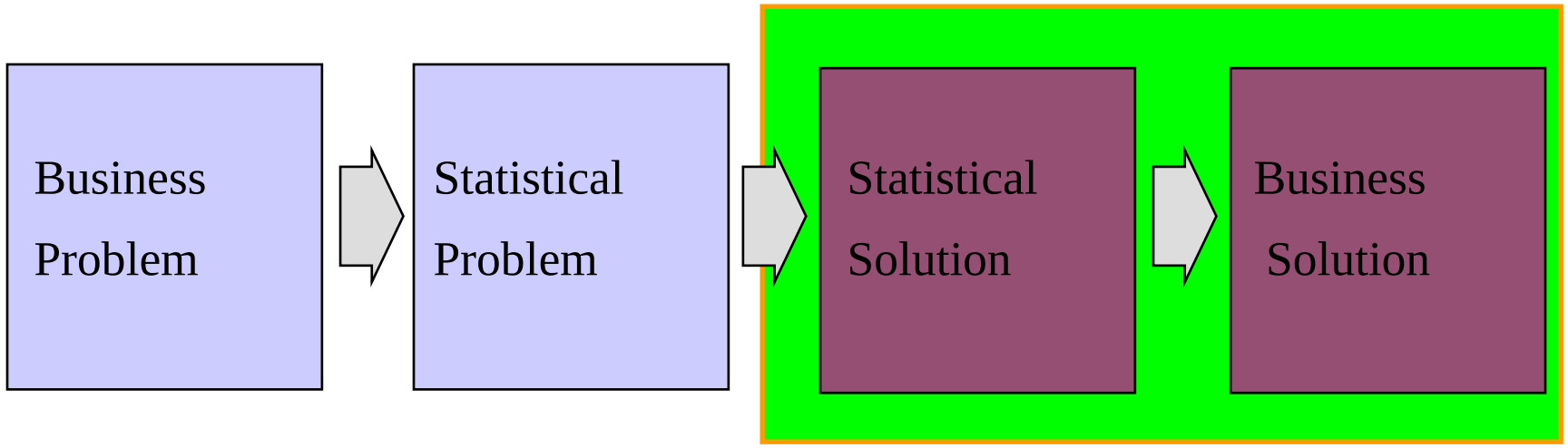
See Appendix for Case Studies

6σ

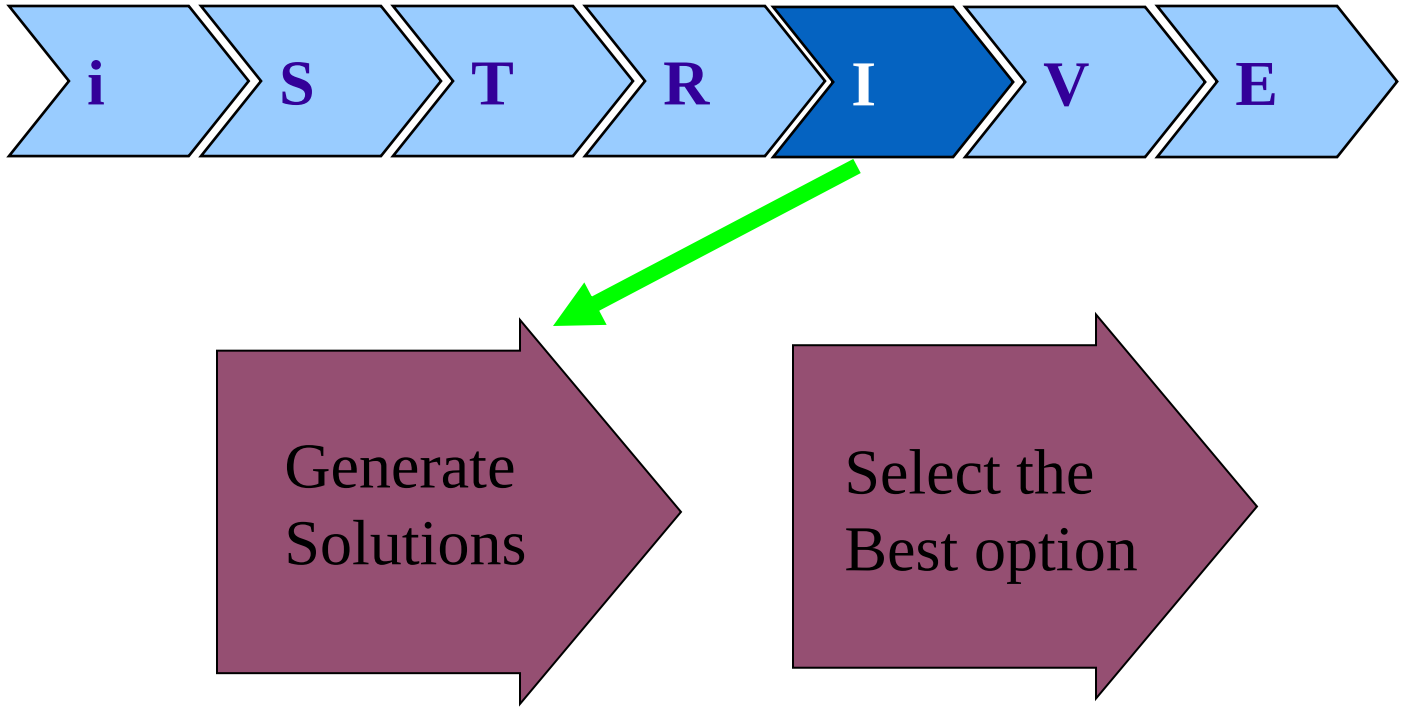
Green Belt Program

IMPROVE

Module 6

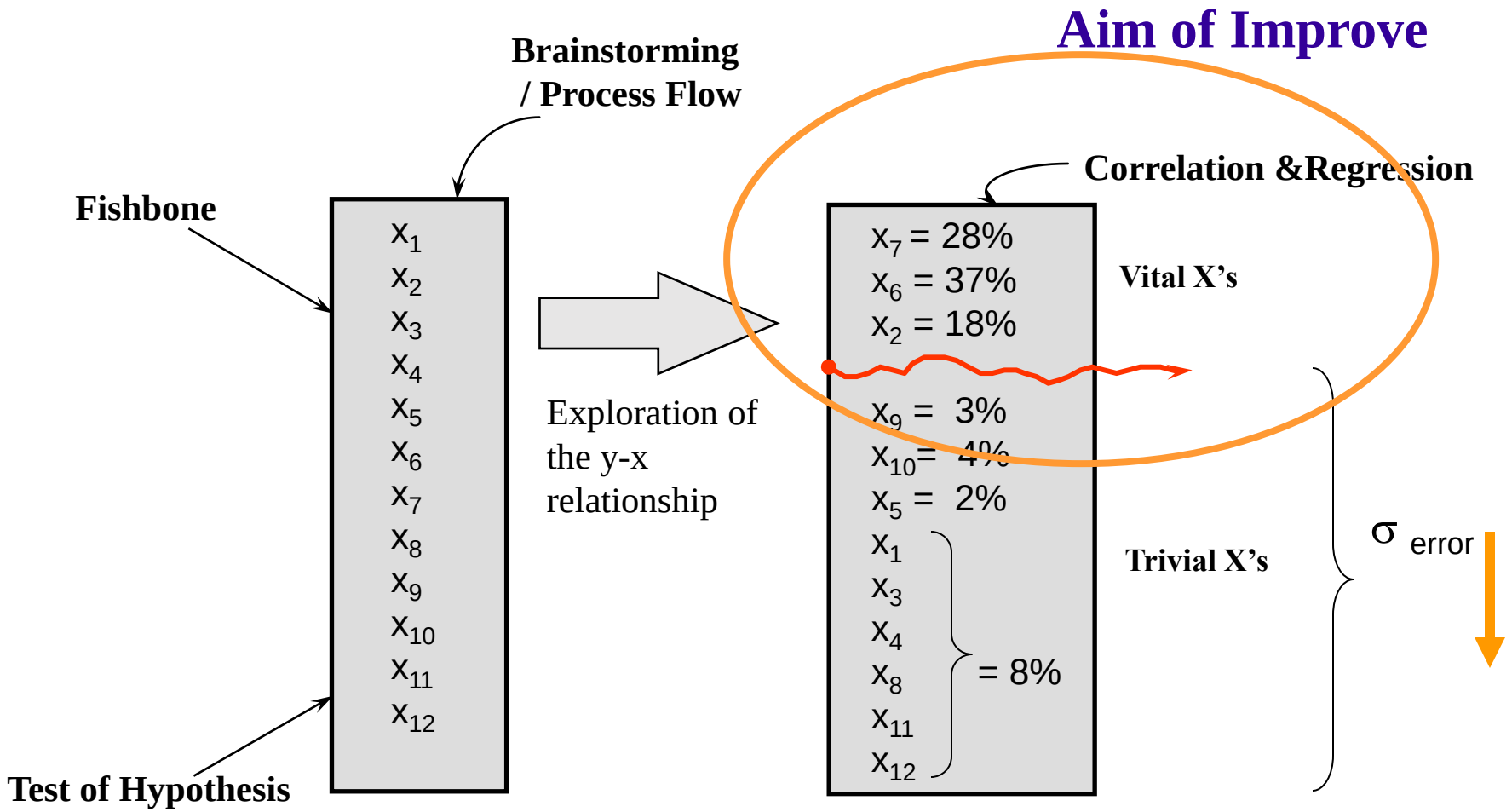


Develop feasible Solutions to improve performance of y



- Develop “feasible solutions” to improve y
- Determine the Operating limits of x’s

- Analyze the Cost benefit aspects of the ‘feasible solutions’
- Consider alternate technologies/ solutions (if required)
- Develop a Pilot Plan to validate the selected Solution


Aim of Improve
Brainstorming
/ Process Flow
Fishbone
Correlation &Regression
x1
x2
x3
x4
x5
x6
x7
x8
x9
x10
x11
x12
Exploration of
the y-x
relationship
x7 = 28%
x6 = 37%
x2 = 18%
Vital X's
x9 = 3%
x10 = 4%
x5 = 2%
x1
x3
x4
x8
x11
x12
= 8%
Trivial X's
σ error
Test of Hypothesis

Validated x's / AICS Delight Requirements	x_1	x_2	.....			x_n
Requirement 1			○			△
Requirement 2		△			○	
Requirement 3	○		◉			
				◉		
		△				
Requirement 4						◉
Total						

Vital x's

◉ = 9 Strong Relationship

○ = 3 Medium Relationship

△ = 1 Weak Relationship

Attributes of Project y

Validated x's / AICS Delight Requirements	Milk Purity	Coffee Powder Quality	Heating Process	Mixing Ratio	Cup Cleaning Process
Hot Coffee			◉		
Clean Cup					◉
No Excess Sugar		△		◉	
Strong	○	◉		○	
Total	3	10	9	12	9

Vital x's

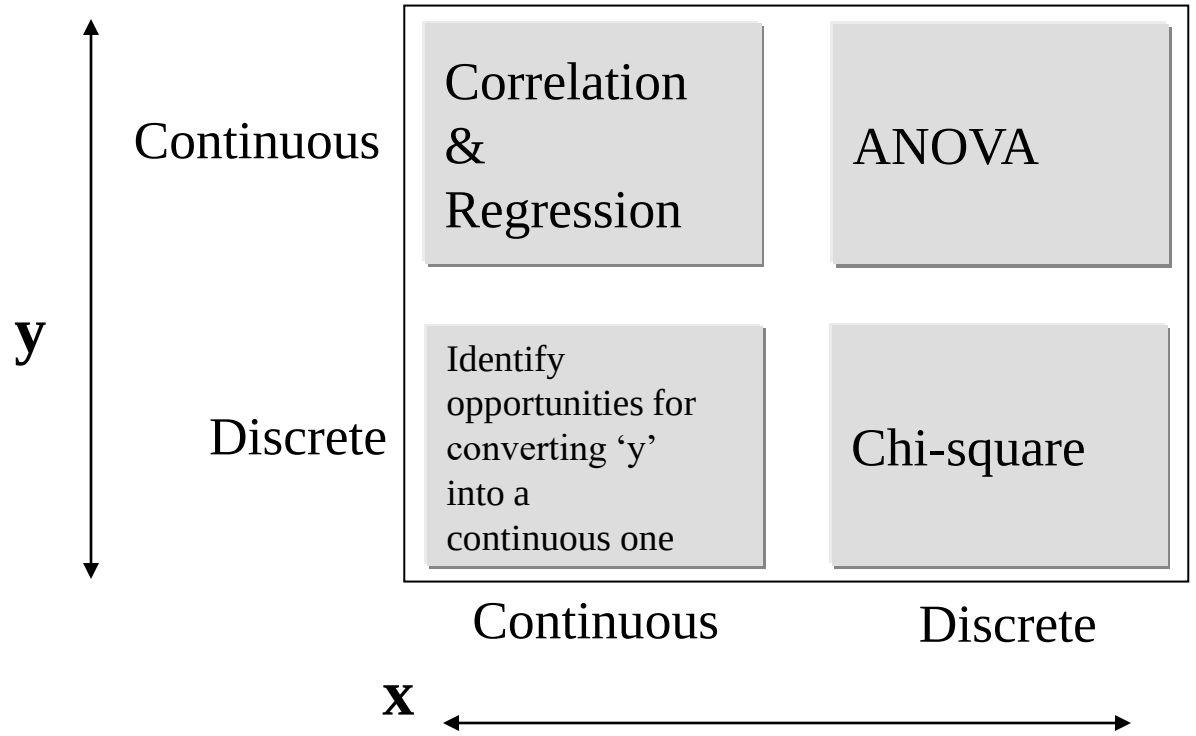
◉ = 9 Strong Relationship

○ = 3 Medium Relationship

△ = 1 Weak Relationship

Correlation and Regression

- To discover whether a substantial relationship exists between ‘y’ & a particular ‘x’
- Depending upon the data characteristics of y & x, we can choose the appropriate tool



- If we want to associate ‘y’ with a ***single*** ‘x’, we can use correlation

- Correlation is about predicting the movement in values in ‘y’ when ‘x’ changes

- Statistical significance of that movement is denoted by correlation coefficient ‘r’

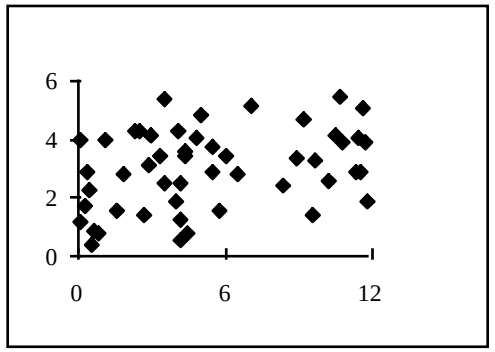
- ‘r’ is always between –1 & +1

    - Positive value of ‘r’ means direction of movement in both variables is same
    - Negative value of ‘r’ means direction of movement in both variables is inverse
    - Zero value of ‘r’ means no correlation between the two variables

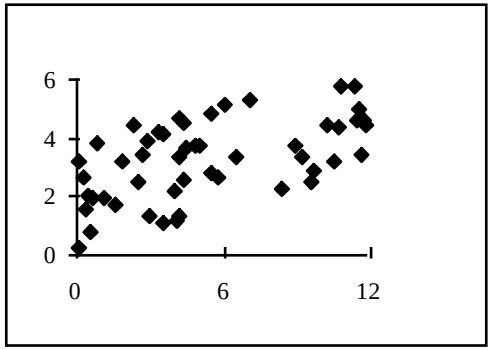
- Higher the value of ‘r’, stronger the correlation between ‘y’ & ‘x’

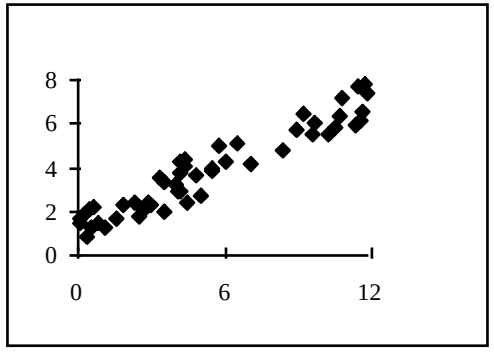
- An ‘r’ value of > + 0.85 or < - 0.85 indicates a strong correlation

r = 0.05

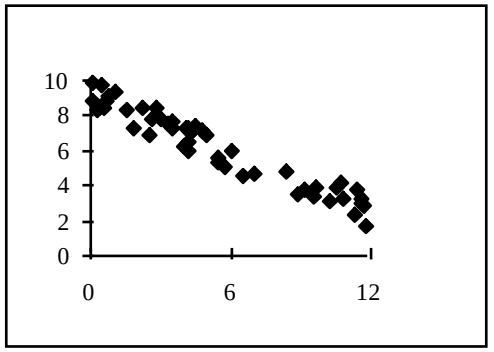


r = 0.50





r = 0.95



r = –0.95

KEY CONCEPTS

- r = 0 indicates no linear relationship.
- r > 0.85 or r < –0.85 indicates a reasonably strong relationship.
- A negative correlation provides as much strength of relationship as a positive one.
- Correlations of 0.5 and below are hard to see.
- Correlation only measures the relationship between the output (y) and one input variable (x).

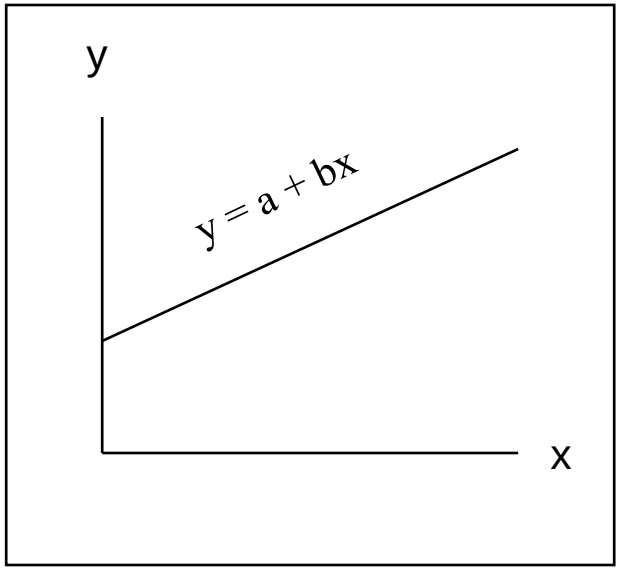
- Correlation tells us only about the direction of movement
- Correlation does not indicate *degree of movement* in one variable with respect to movement in another
- Regression of ‘y’ on ‘x’ results in a transfer function equation that can be used to *predict* the value of ‘y’ for given values of ‘x’

$$y = \mathbf{f}(x)$$

- ‘y’ can be regressed on one or more x’s simultaneously
  - Simple linear regression is for one x
  - Multiple linear regression is for more than one x’s

What?

A mathematical means of describing a relationship between the “y” and the “x”s - creating a “model” of the process.



y = a + bx ; Where,
y = Dependent variable / output / response
x = Independent variable / input / predictor
a is the y intercept
b is the slope of the line/regression coefficient

Why?

- To find the potential Vital Few “x”s
- To predict / forecast the “y”
- To optimize the “y”
- To determine where to set the “x”s to optimize “y”

r: The correlation coefficient (r) for multiple regression. The closer to +/ - 1, the better the fit of the model. ‘0’ indicates no linear relationship.

R-Sq: The correlation coefficient squared (R^2). A value of R^2 closer to 100% indicates that there is a possible relationship, and more variation is explained.

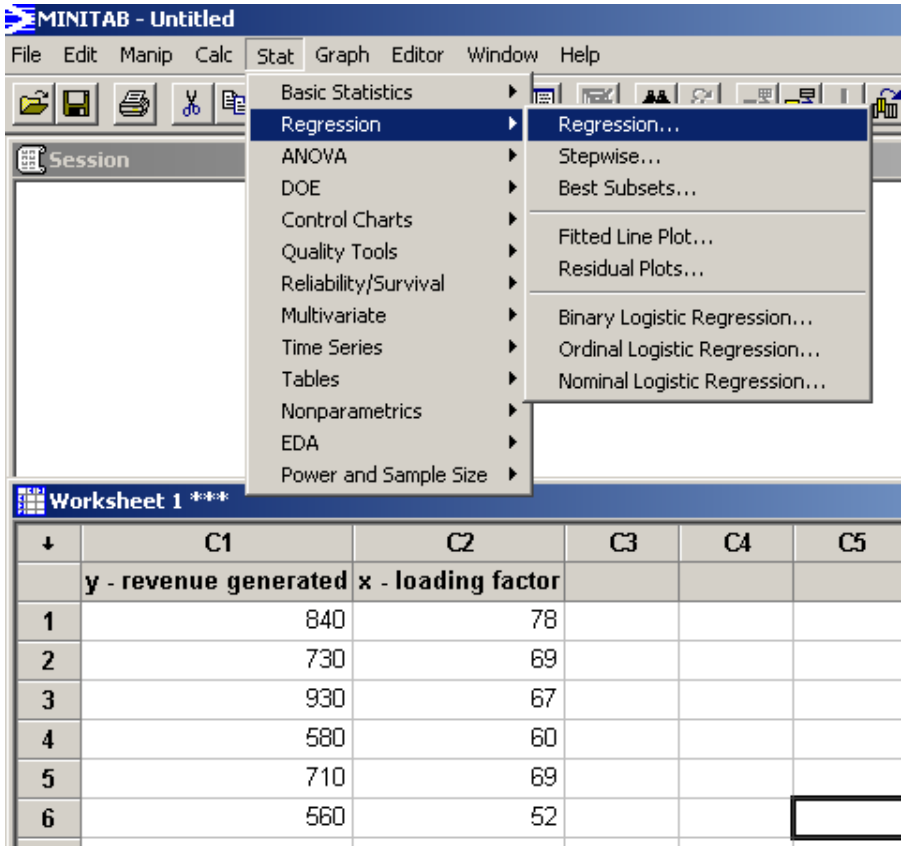
R-Sq (Adj): Adjustment of R^2 for an overfit condition. (Takes into account the number of terms in the model).

Standard Error of the Estimate (s) Expected deviation of data about the predictive “surface” $s = (MS_{error})^{1/2}$

- Suppose that Corporate wants to study the relationship between the Loading factors and the revenue generated across different units on a regular basis. It collects the following data for last one years. The purpose is to forecast the revenue for the next year and determine right size of the units.

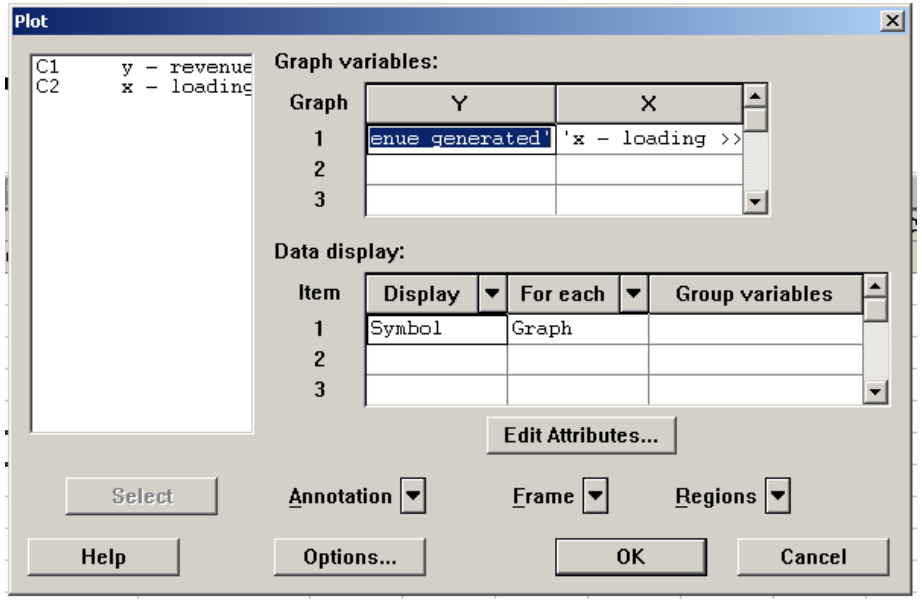
Month	Revenue (in millions)	% Loading Factor
1	840	78
2	730	69
3	930	67
4	580	60
5	710	69
6	560	52
7	820	63
8	880	84
9	900	72
10	440	41
11	550	48
12	930	77

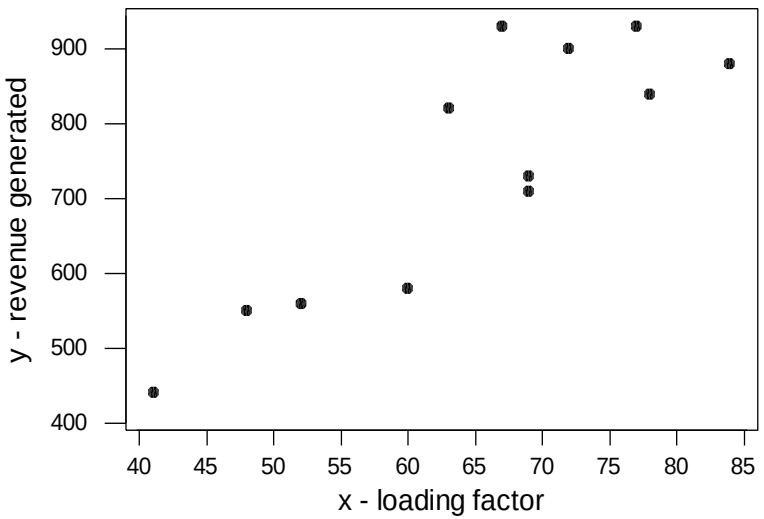
Minitab output



	C1	C2	C3	C4	C5
	y - revenue generated	x - loading factor			
1	840	78			
2	730	69			
3	930	67			
4	580	60			
5	710	69			
6	560	52			

Graph > Plot





Minitab output

Regression Analysis: y - revenue generated versus x - loading factor

The regression equation is
y - revenue generated = -10 + 11.5 x - loading factor

Predictor	Coef	SE Coef	T	P
Constant	-9.7	136.1	-0.07	0.945
x - load	11.520	2.057	5.60	0.000

Fail to reject H_o

Accept H_a

S = 87.80 R-Sq = 75.8% R-Sq(adj) = 73.4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	241811	241811	31.37	0.000
Residual Error	10	77080	7708		
Total	11	318892			

- P-value of 0.000 for x – loading factor denotes that revenue is statistically significant to have a *non-zero* regression coefficient
- R-Sq = 75.8 % = r^2. This is called the *coefficient of determination*
- Also, R-Sq = $SS_{Regression}$ / SS_{Total}
- The R-Squared value is the proportion of variability in the y variable accounted for by the predictors. In other words, 75.8% of variation in 'y' is explained by 'x' in the fitted model

p-value of the Constant

H_o: The line passes through the origin (0,0)... (0 loading factor = 0 revenue)

H_a: The line does not pass through the origin (0,0)..(0 loading factor = 0 revenue)

p-value of the "x" variable

H_o: Slope = 0

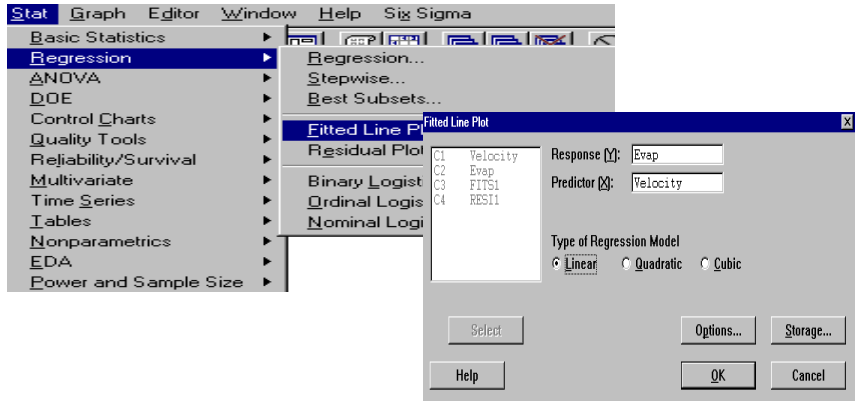
H_a: Slope = 0

or, another way of saying it:

H_o: The "x" is not significant

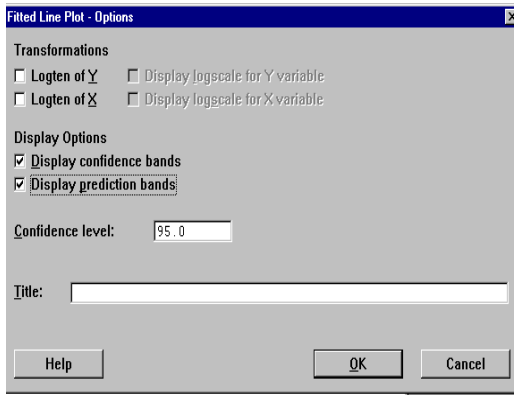
H_a: The "x" is significant

Stat>Regression>Fitted Line Plot



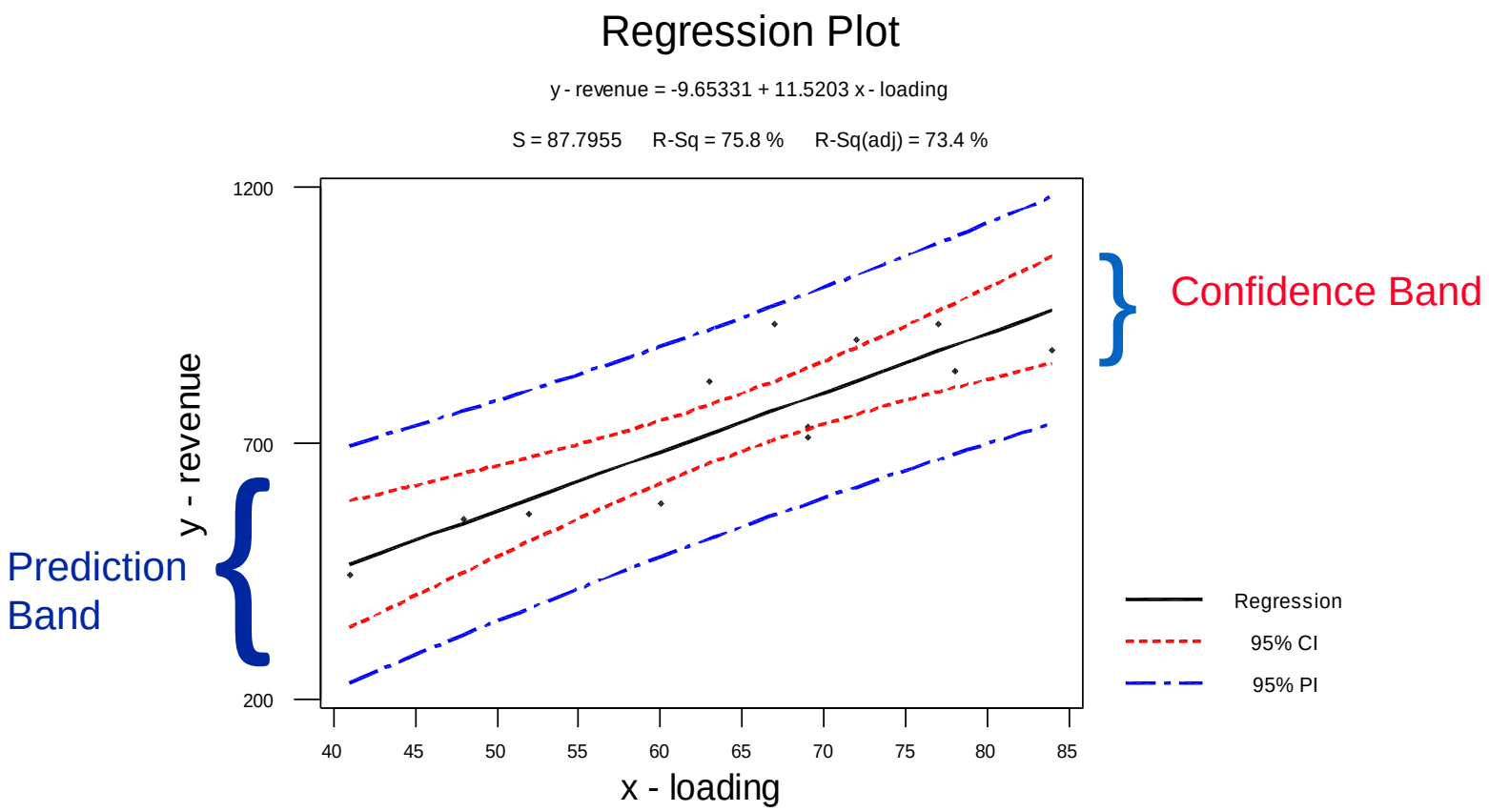
Click on “Options”

Click these Options to display more information in the graphical output



Click “OK” twice

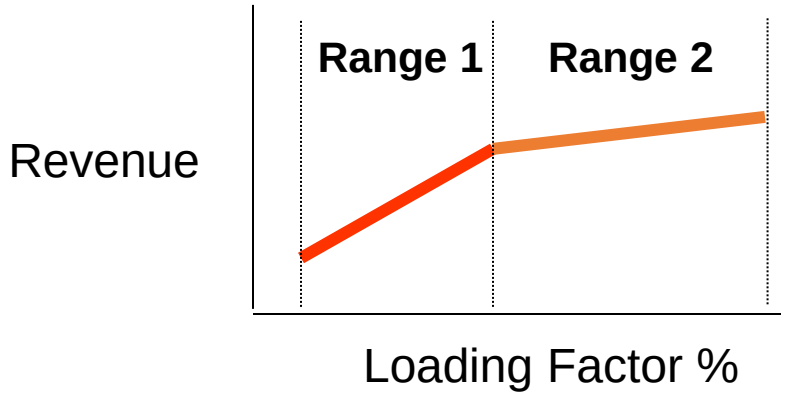
- “Fitted Line Plot” provides:
  - Regression Analysis in the Session Window
  - A plot showing the Least Squares fit for the line *
  - A plot showing Confidence Intervals (C.I.) and Prediction Intervals (P.I.)



C.I. = Confidence Interval (95% confidence that the means of all data will fall within this band)

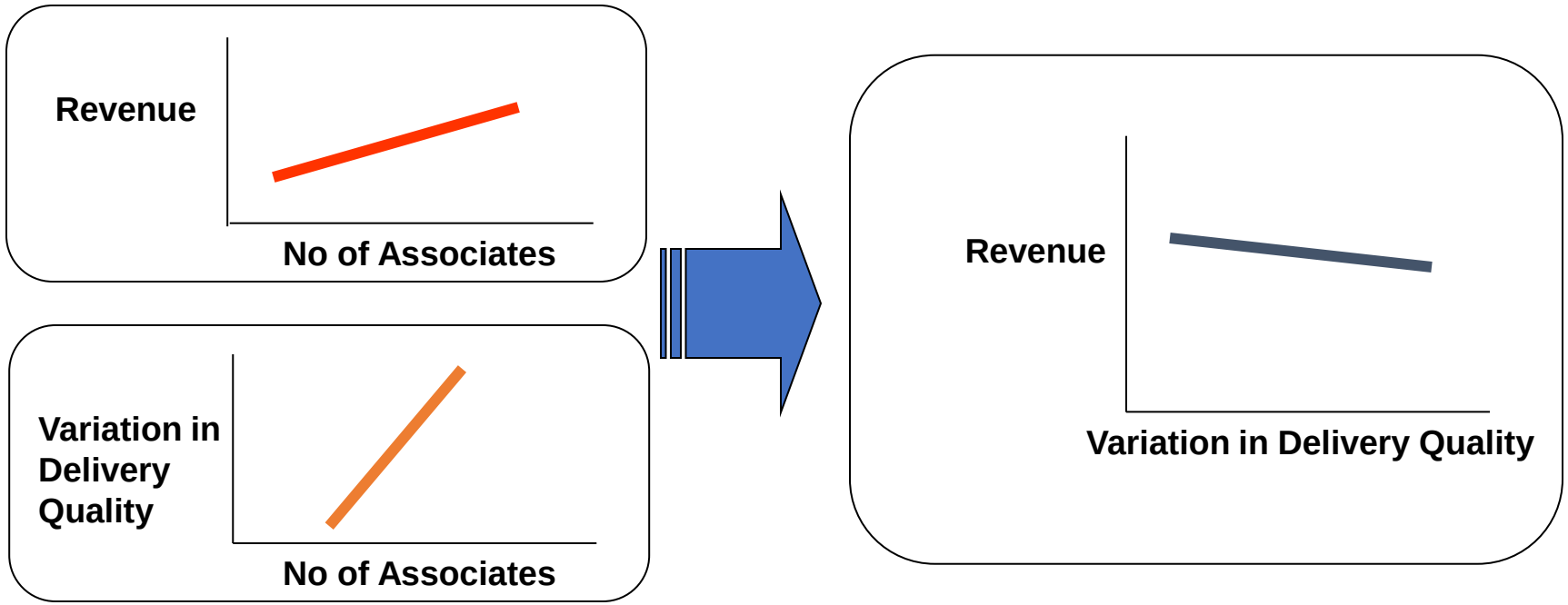
P.I. = Prediction Interval (95% confidence that the individual data points will fall within this band)

- Don’t extrapolate beyond the range
  - Project teams get tempted to extrapolate the results beyond the range of collected data. In one of the previous examples, higher loading factors may not see same increase in revenue, & hence the regression equation between revenue & loading factor may change



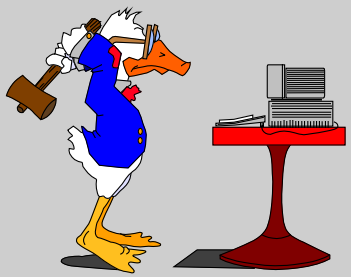
- Don’t assume causation
  - Regression equation denotes a relationship only. This in no way means that a change in one variable *causes* change in another. If variation in Delivery Quality and Revenue generated in the organisation may be inversely related, but no causation. The increase in both factors could be due to third factor – increase in number of associates

  - In other words, both of them may be dependent variables themselves



- **Draw the Improved Process Map after**
  - **establishing the relationship between “y” and the root causes “x”**
- **Take the improved process through “FMEA” to identify Potential Failure Modes**
- **Develop appropriate Controls**

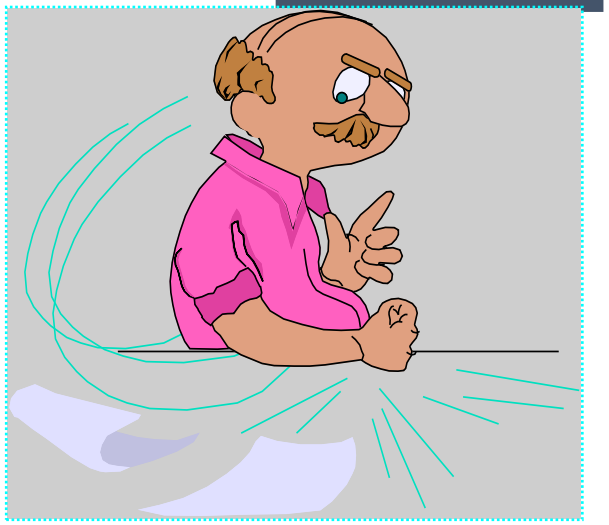
Failure Mode (Defect)



Material or process input

Effect

Controls



Cause

Failure Mode - The way in which a specific process input fails - can be associated with a defect (attribute data) or an out of spec.(continuous data)

Effect - The impact the failure has on the Critical Quality Parameters

Cause - The source of variation that caused the process to fail

Current Controls - Systemized devices in place to prevent or detect the failure

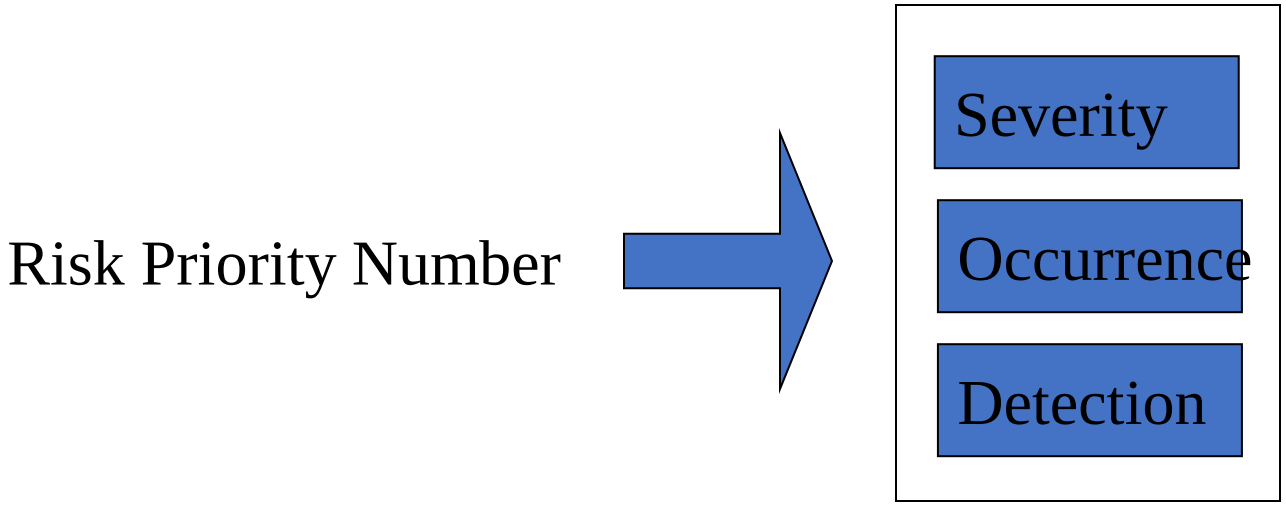
Severity - Importance of an effect on the Critical Quality parameter (1- not severe; 10 - very severe)

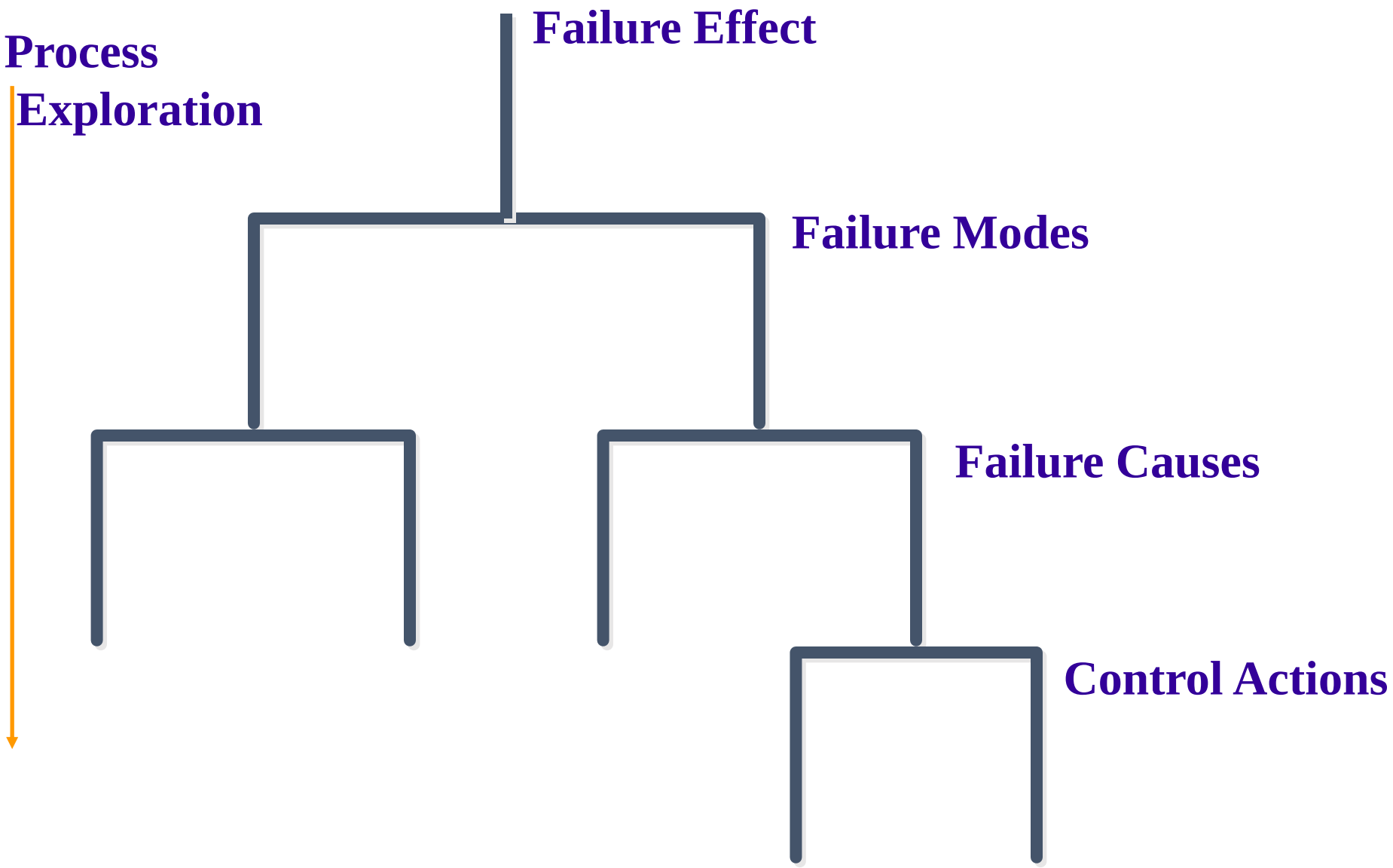
Occurrence - Frequency with which a cause occurs(1 - not likely; 10 - very likely)

Detection - Ability of current control to detect the cause before creating a failure mode (1 - likely to detect; 10 - not likely to detect)

Risk Priority Number

- A numerical calculation of the relative risk of a particular Failure Mode.
- This number is then used to place priority on which items need additional quality planning.




Process
Exploration
Failure Effect
Failure Modes
Failure Causes
Control Actions

Action Plan Matrix						
Root Cause	Counter Measure	Method	Feasibility	Effectiveness	f*e	Action
			f	e		

- Used to plan action for elimination of root causes

- Develop the optimum process parameter combination improving the ‘y’ performance
  - Set
    - vital x’s are at the optimum level
    - the trivial x’s are at the most favorable level
  - Verify
    - the predicted level of ‘y’
    - the noise/error in the process
  - Confirm
    - the new Z-value of the process (performance)

Questions to be answered

1. What is the root cause (validated vital x)?

2. Are there multiple causes for the problem (or multiple x's affecting the Project y's)

3. How can we verify the root cause(s)?

4. Have we established the relationship between critical x's and y's ?

5. What could be done about each cause ?

Deliverables - Phase level

1. Confirmed Vital x's

2. Optimized solutions

3. Pilot plan

Tools

Prioritization/Pugh Matrix **Experimentation** **QFD**

Do's

1. Verify the cause-effect relationships. Eliminate the possibility of spurious correlation's

2. Quantify the main effects causing variation in Project y's

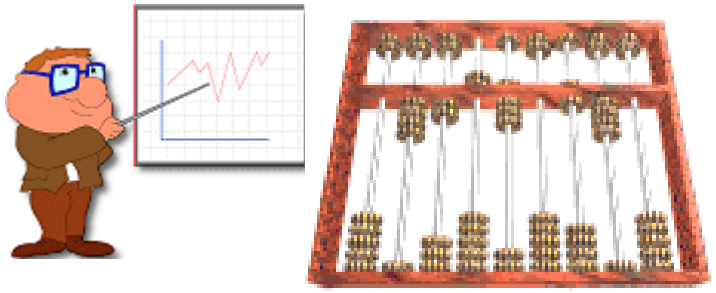
2. To evaluate solutions, identify the 'musts' and the 'wants'. Solutions that pass the 'must' criteria can be evaluated to identify the best solution - or to combine elements of solutions to for a better one

Don'ts

1. Identify one solution to a problem.

2. Decide on "a solution" based on single criteria viz. Economy /Cost Savings

Proposed Solution(s)

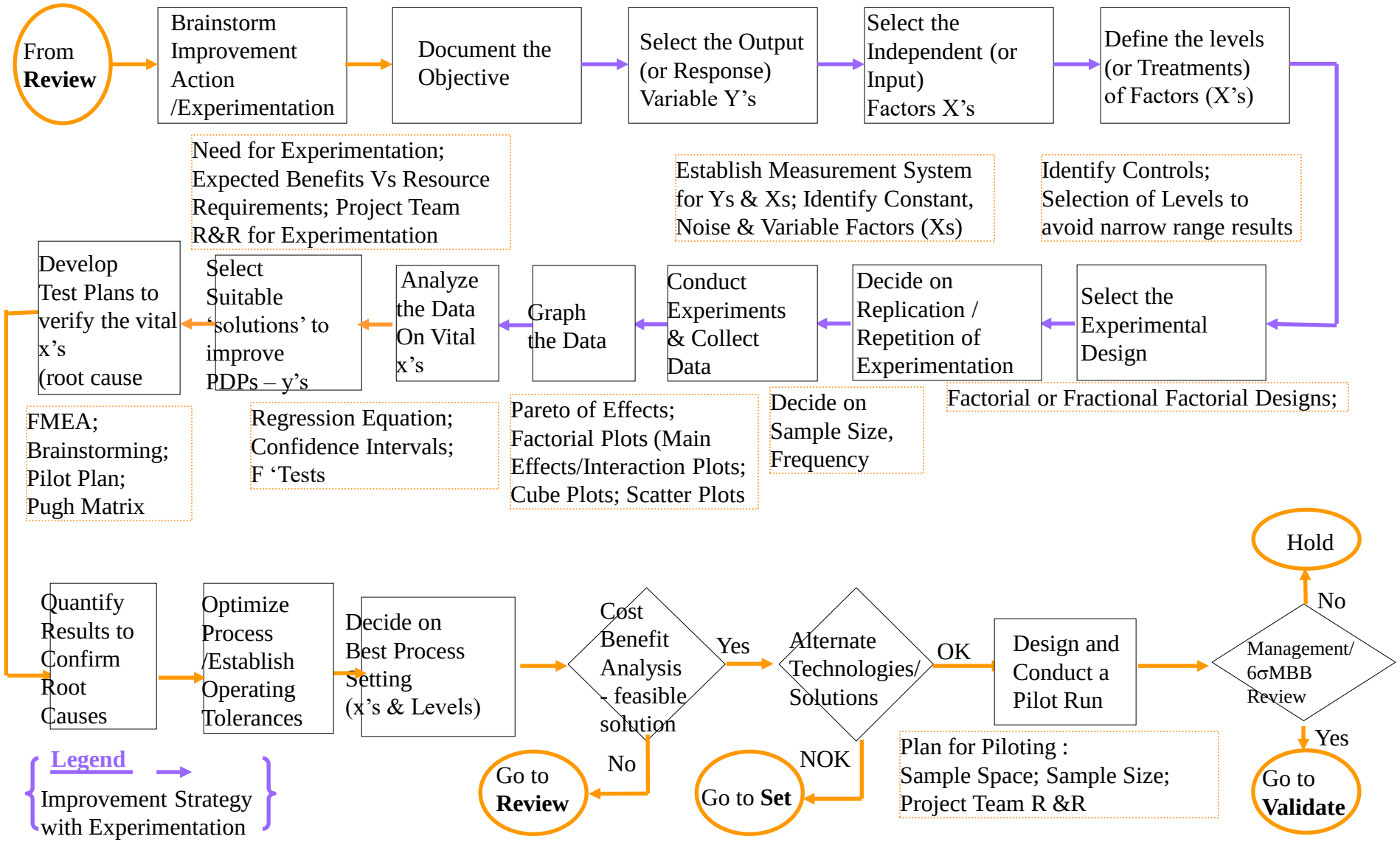


Benefit-Cost Analysis

Select the best solution

Pilot the Solution

Improve the process performance to ‘optimal’ levels


From **Review**
Brainstorm Improvement Action /Experimentation
Document the Objective
Select the Output (or Response) Variable Y's
Select the Independent (or Input) Factors X's
Define the levels (or Treatments) of Factors (X's)
Need for Experimentation; Expected Benefits Vs Resource Requirements; Project Team R&R for Experimentation
Establish Measurement System for Ys & Xs; Identify Constant, Noise & Variable Factors (Xs)
Identify Controls; Selection of Levels to avoid narrow range results
Develop Test Plans to verify the vital x's (root cause
Select Suitable 'solutions' to improve PDPs – y's
Analyze the Data On Vital x's
Graph the Data
Conduct Experiments & Collect Data
Decide on Replication / Repetition of Experimentation
Select the Experimental Design
FMEA; Brainstorming; Pilot Plan; Pugh Matrix
Regression Equation; Confidence Intervals; F 'Tests
Pareto of Effects; Factorial Plots (Main Effects/Interaction Plots; Cube Plots; Scatter Plots
Decide on Sample Size, Frequency
Factorial or Fractional Factorial Designs;
Quantify Results to Confirm Root Causes
Optimize Process /Establish Operating Tolerances
Decide on Best Process Setting (x's & Levels)
Cost Benefit Analysis - feasible solution
Yes
No
Go to **Review**
Alternate Technologies/ Solutions
OK
NOK
Go to **Set**
Design and Conduct a Pilot Run
Plan for Piloting : Sample Space; Sample Size; Project Team R &R
Management/ 6σMBB Review
No
Hold
Yes
Go to **Validate**
Legend
Improvement Strategy with Experimentation

Microsoft Word
Document

(A Case of Proximity)

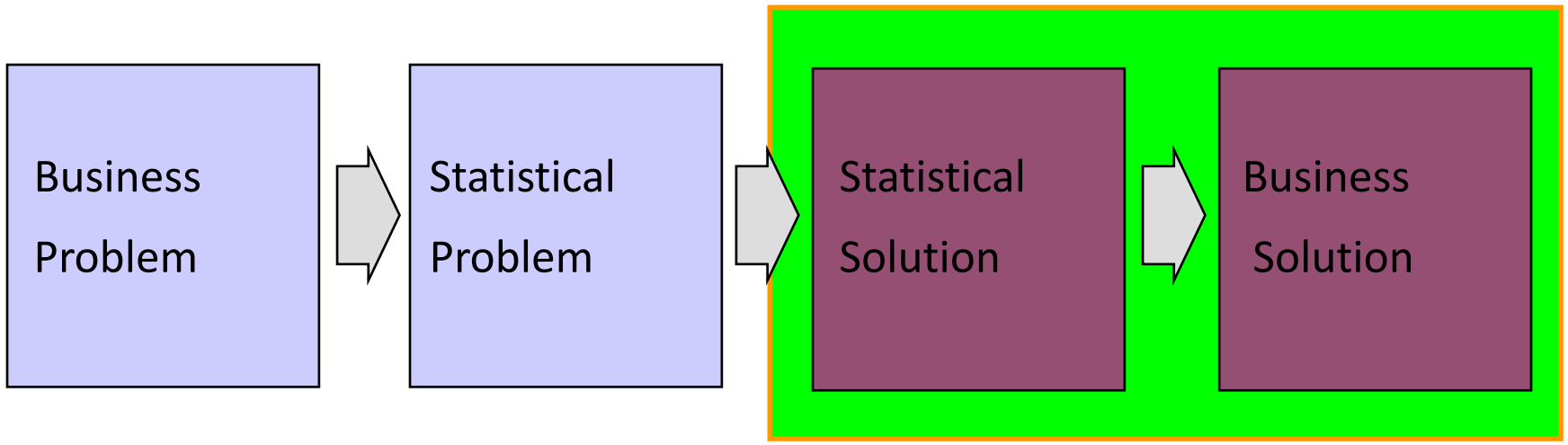
See Appendix for Case Studies

6σ

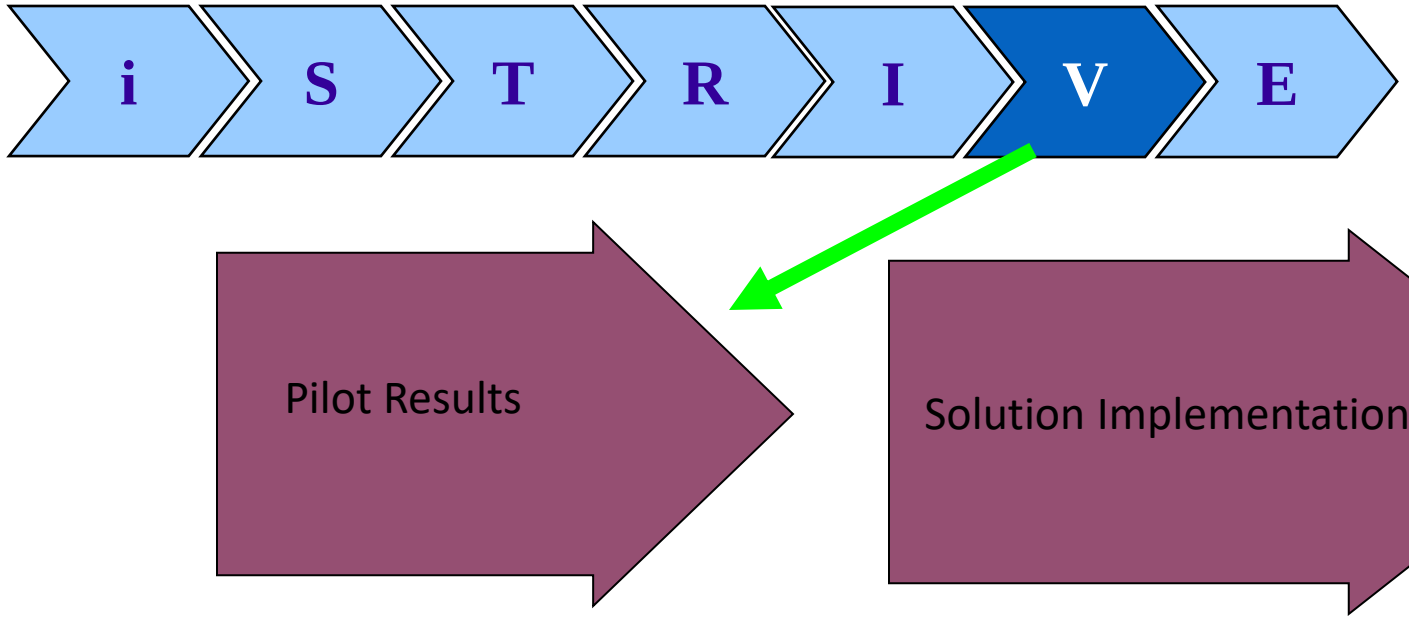
Green Belt Program

VALIDATE

Module 7



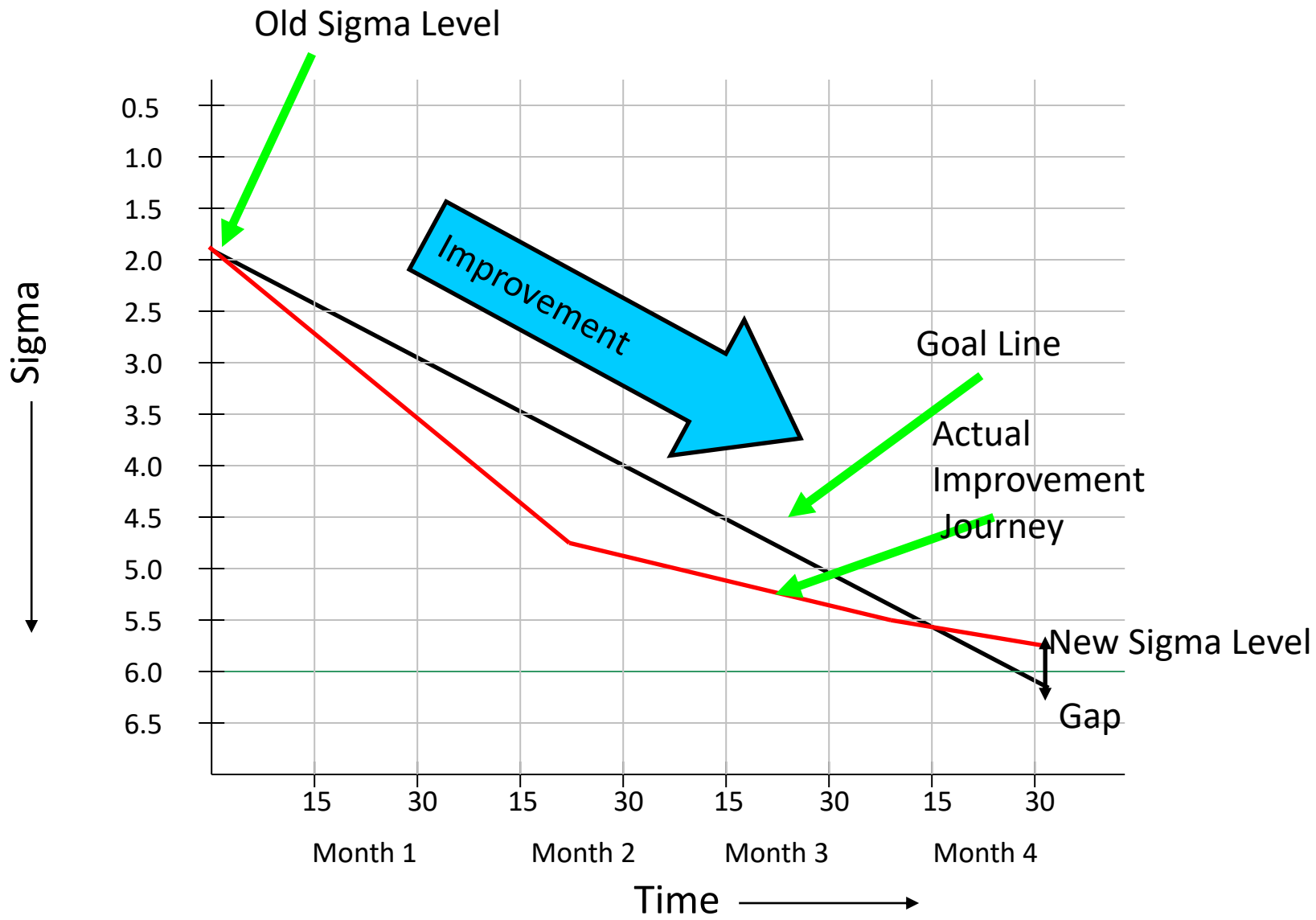
- Pilot run the Modified Process Flow to validate results
- Develop Process control system to sustain improvements



- Prepare the modified process flow as per new solution
- Conduct Pilot run to confirm the performance results of y with set targets [New Z value]
- Study the influence of 'solution' on other process measures (y's) and on Circle Measure (Y)
- Evaluate the 'financial gains'
- Assess the Risk involved in Solution implementation

- Develop a Risk Mitigation Plan for solution implementation
- Develop a Change Management Strategy for solution implementation
- Develop a Process Control System for sustenance of new level of performance (mistake-proofing)
- Develop Roll-out Plan and Implement Solution.

Validate the Process Improvements


Old Sigma Level
Sigma
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
Improvement
Goal Line
Actual
Improvement
Journey
New Sigma Level
Gap
15
30
15
30
15
30
15
30
Month 1
Month 2
Month 3
Month 4
Time

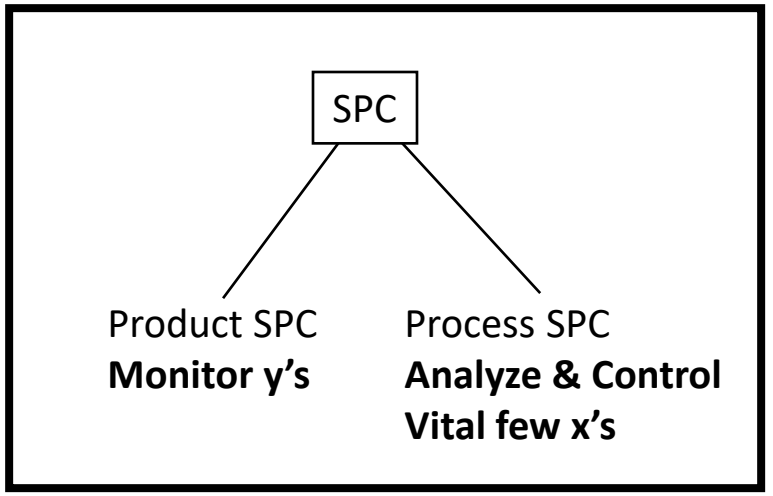
Statistical Process Control (SPC)

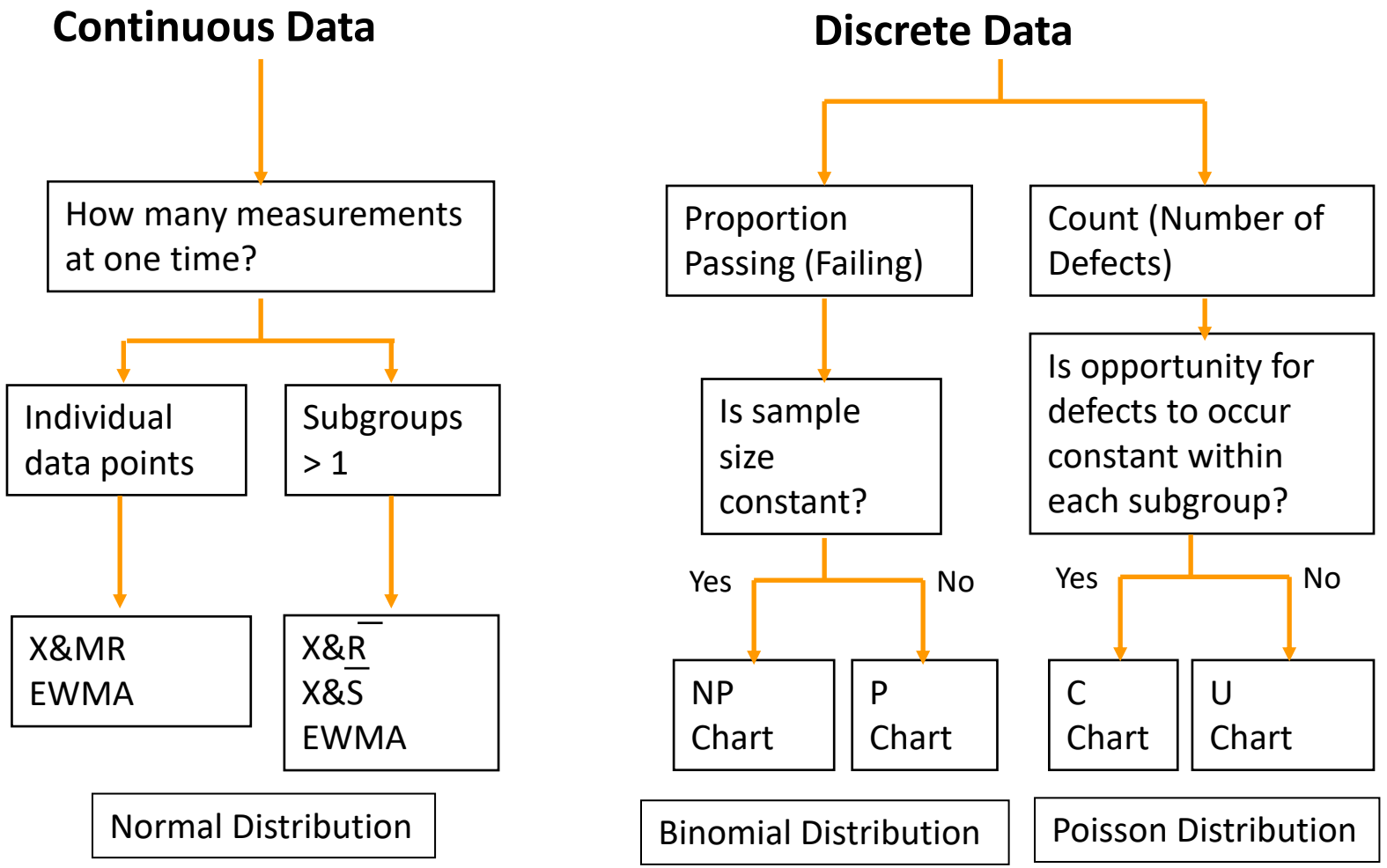
Controlling the x’s

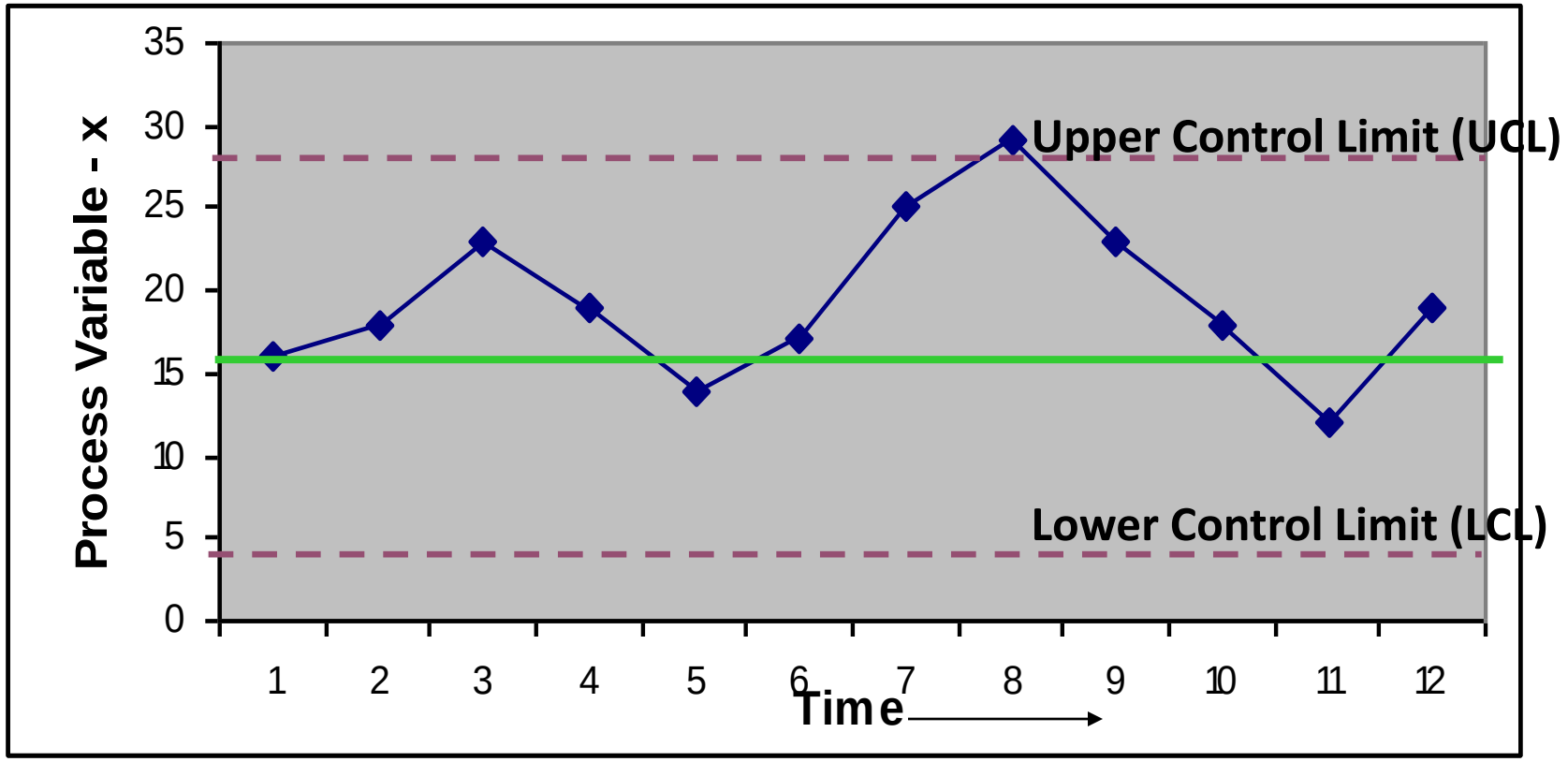
- SPC was developed by Walter Shewhart in 1924.
- Traditionally SPC has been used to monitor and control the output parameters of the process (y's)

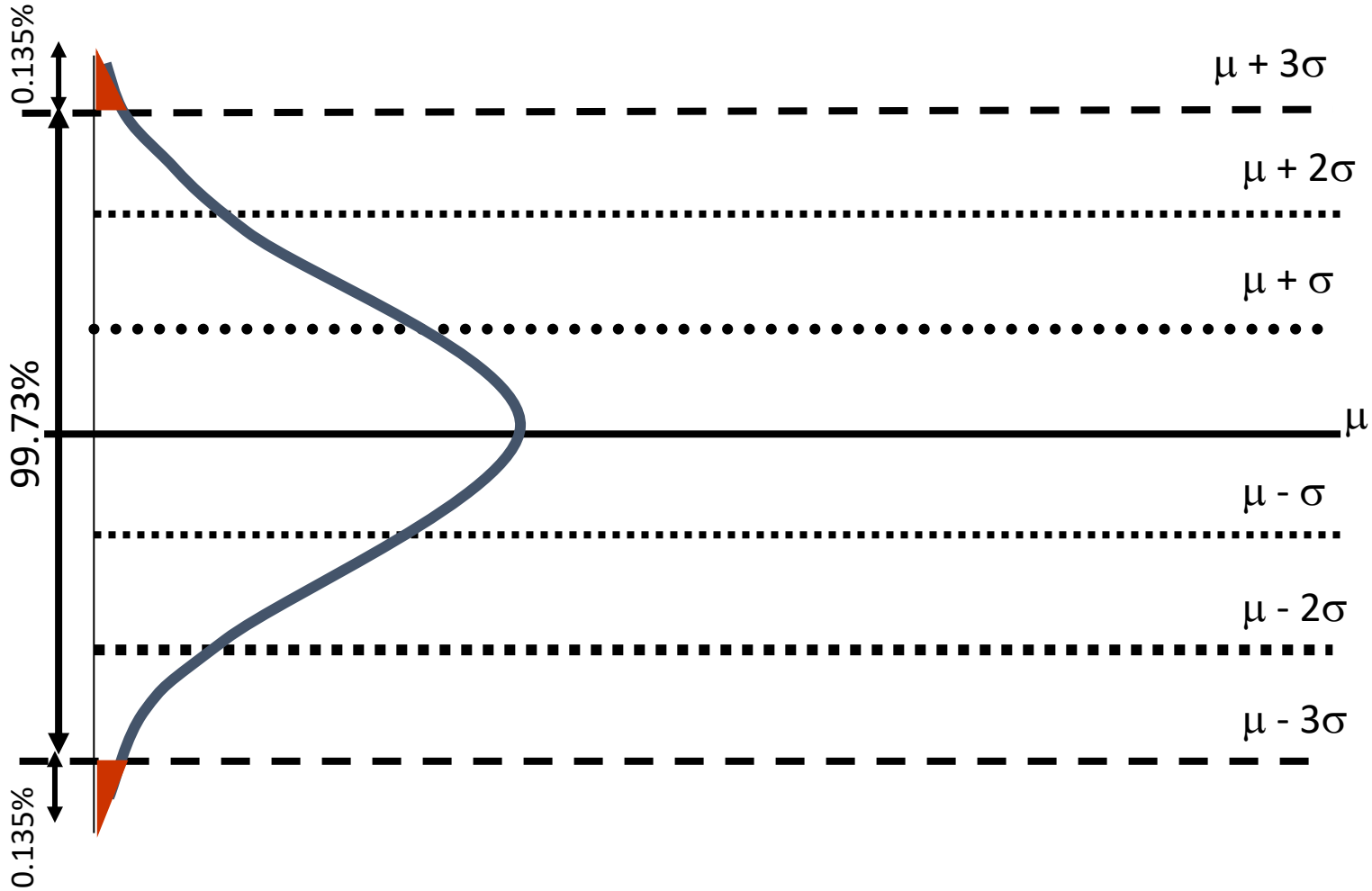
SPC is used to

- Analyze & control process performance
- Proactively control processes
- Distinguish between natural & assignable variation
- Identify & Prevent process from special causes

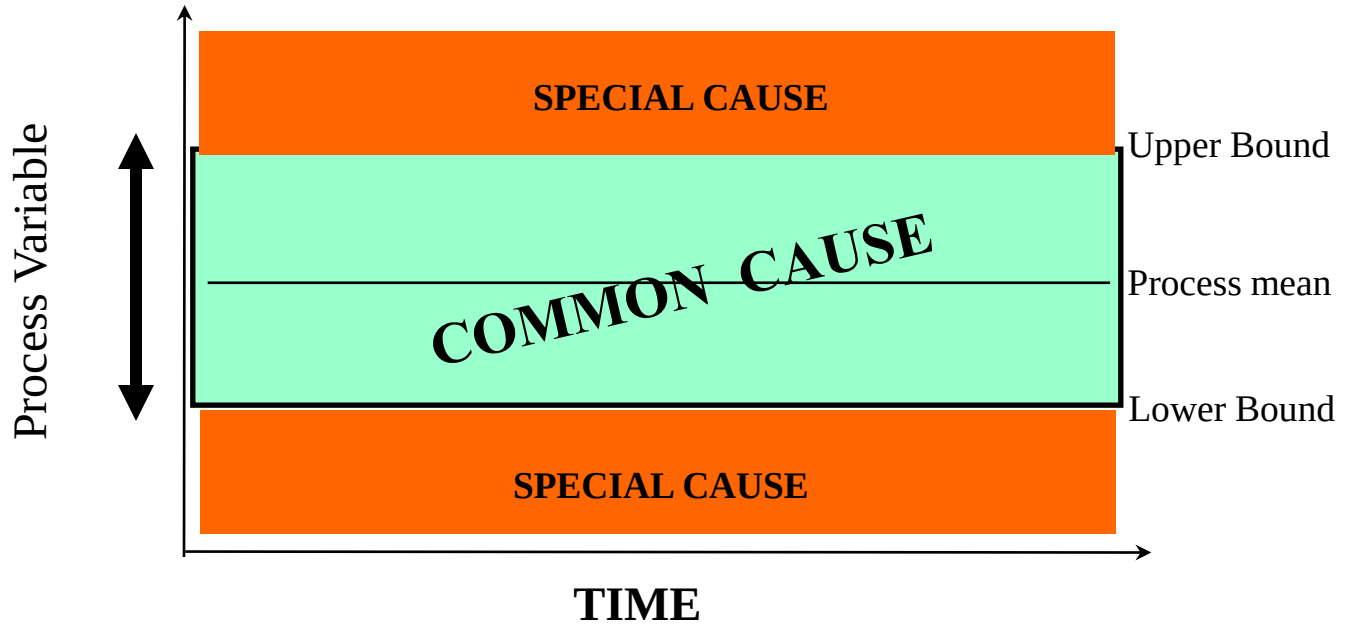



Continuous Data
How many measurements at one time?
Individual data points
Subgroups > 1
X&MR
EWMA
X&R̄
X&S
EWMA
Normal Distribution
Discrete Data
Proportion Passing (Failing)
Count (Number of Defects)
Is sample size constant?
Is opportunity for defects to occur constant within each subgroup?
Yes
No
Yes
No
NP Chart
P Chart
C Chart
U Chart
Binomial Distribution
Poisson Distribution


Process Variable - x
35
30
25
20
15
10
5
0
Upper Control Limit (UCL)
Lower Control Limit (LCL)
1
2
3
4
5
6
7
8
9
10
11
12
Time



A Standard control chart uses control limits at three standard deviations from the data mean.



Common Causes

Present all the time

Have a small effect individually

Result in random variation

Effects we will tolerate

Special Causes

Not always present

Come from outside influences

Typically have a bigger individual influence than common causes

Effects we want to hear about

- **Monitor the process for assignable causes and eliminate them**
- **Limit variability to Inherent Variability**

Benefits -

- Process variability is restricted
- Process is repeatable
- Quality of future process output(s) are predictable

- Process is said to be under “Statistical Control” if the process is not being affected by the Special Causes

- All the points must be within the control limits randomly dispersed about the average line for an in-control system.

- Control doesn’t necessarily mean that the Product or Service is meeting the Needs.

- Control means - Process is consistent; Process Variation is predictable**.**

Out-of-Control Signals

- One or more points either above UCL or below LCL
- Run of Two of the Three consecutive points in the same zone (area between UCL/LCL and Central Line divided into three equal zones) or beyond.
- Run of Eight or more points above or below the central line.
- Run of Six or more points up or down

I - Chart

Central Line : Target or $\bar{\bar{X}}$

UCL : $\bar{\bar{X}} + 3\,\overline{MR} / d_2$

LCL : $\bar{\bar{X}} - 3\,\overline{MR} / d_2$

MR = Moving range = $\sigma = \overline{MR} / d_2$;

Process Capability = 6σ

MR - Chart

Central Line : $\overline{MR}$

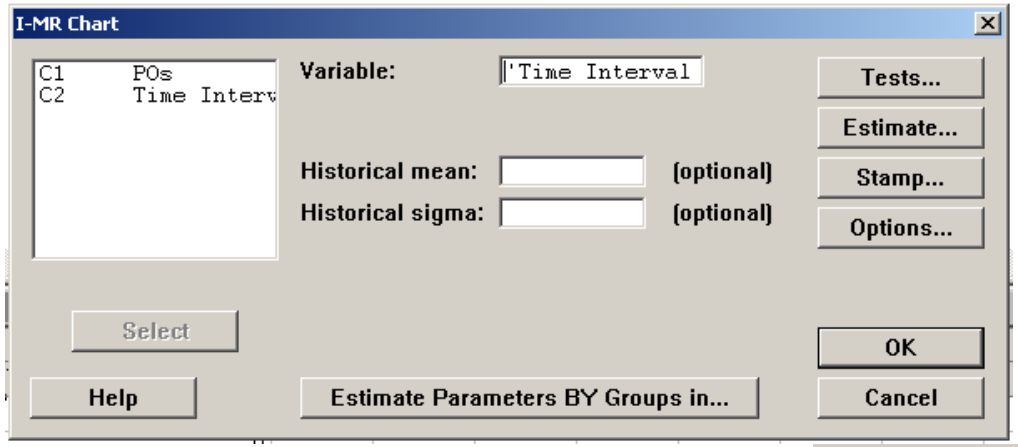
UCL : $D_4\,\overline{MR}$

LCL : $D_3\,\overline{MR}$

D_4 , D_3 & d_2 to read from tables for n = 2

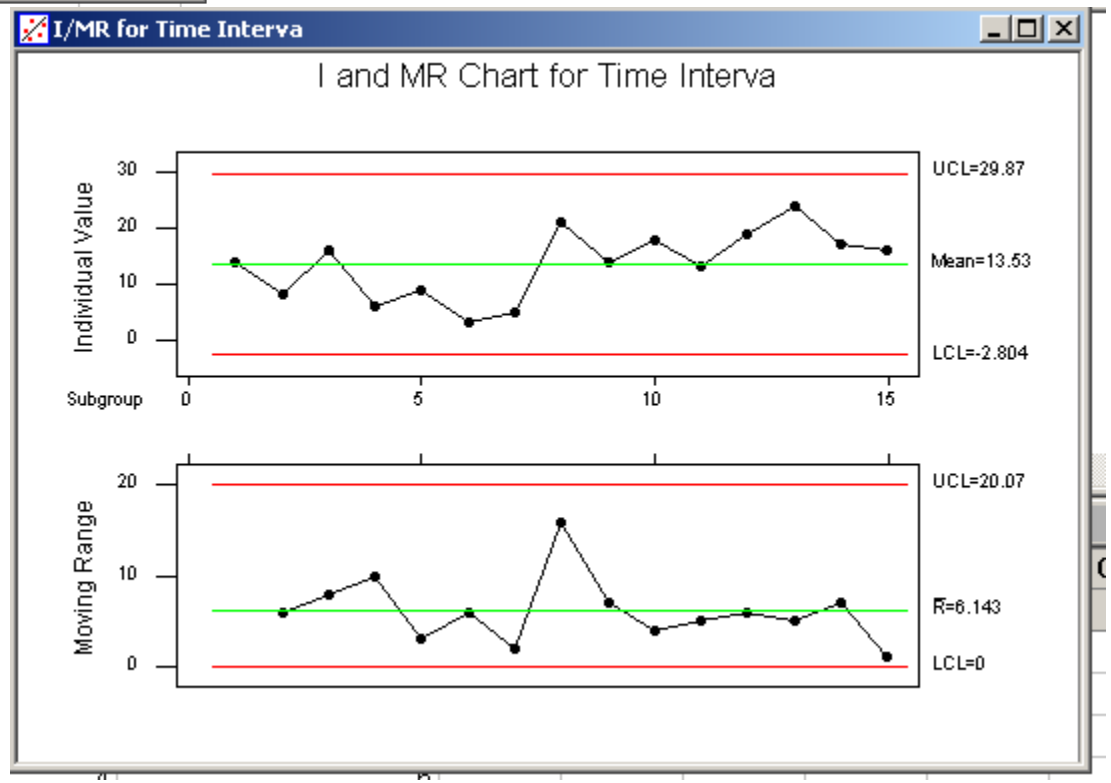
- Time interval between RFP raising to PO release is recorded on a regular basis. The cycle time is a vital ‘x’ to successfully meet the requirements of the indenter. This is monitored periodically to keep it within desired operating range.
- Data on time interval (in days) is collected over 15 cases.

PO's raised	Time Interval for release
1	14
2	8
3	16
4	6
5	9
6	3
7	5
8	21
9	14
10	18
11	13
12	19
13	24
14	17
15	16



- STAT > CONTROL CHARTS > I-MR

- Minitab gives the following output:



- Interpreting Results
  - Mean of the given data points is 13.63 which is the central line of the top chart
  - Standard deviation of the data is 5.445 (S_{ST}) which results in three sigma control limits of UCL = 13.53 + 16.34 = 29.87 & LCL= 13.53 – 16.34 = 2.81
  - Note that S_{ST} can not be calculated from usual standard deviation formula since there are no sub-groups (you can use capability analysis)
  - First point in the lower chart is the absolute difference between first & second data point = I 14 – 8 I = 6
  - Second point in the lower chart is the absolute difference between second & third data point = I 8 – 16 I = -8…….& so on. Mean of all such differential points is 6.143
  - All observations found to be lying within control limits with random pattern & hence some more readings must be taken to observe the trend

c - Chart

Sample Size : same or constant

Central Line : $\bar{c}$

UCL : $\bar{c} + 3\sqrt{\bar{c}}$

LCL : $\bar{c} + 3\sqrt{\bar{c}}$

c = number of defects observed

u-Chart

Sample Size : Not constant

Central Line : $\bar{u}$

UCL : $\bar{u} + 3\sqrt{\bar{u} / n}$

LCL : $\bar{u} - 3\sqrt{\bar{u} / n}$

n = sub-group size

u = number of defects per unit

No of defects observed in payroll processing from 1000 cases sampled each month is presented below:

Sample No.	No. of defects
1	21
2	24
3	16
4	12
5	15
6	5
7	28
8	20
9	31
10	25
11	20
12	24
13	16

Sample No.	No. of defects
14	19
15	10
16	17
17	13
18	22
19	18
20	39
21	30
22	24
23	16
24	19
25	17
26	15

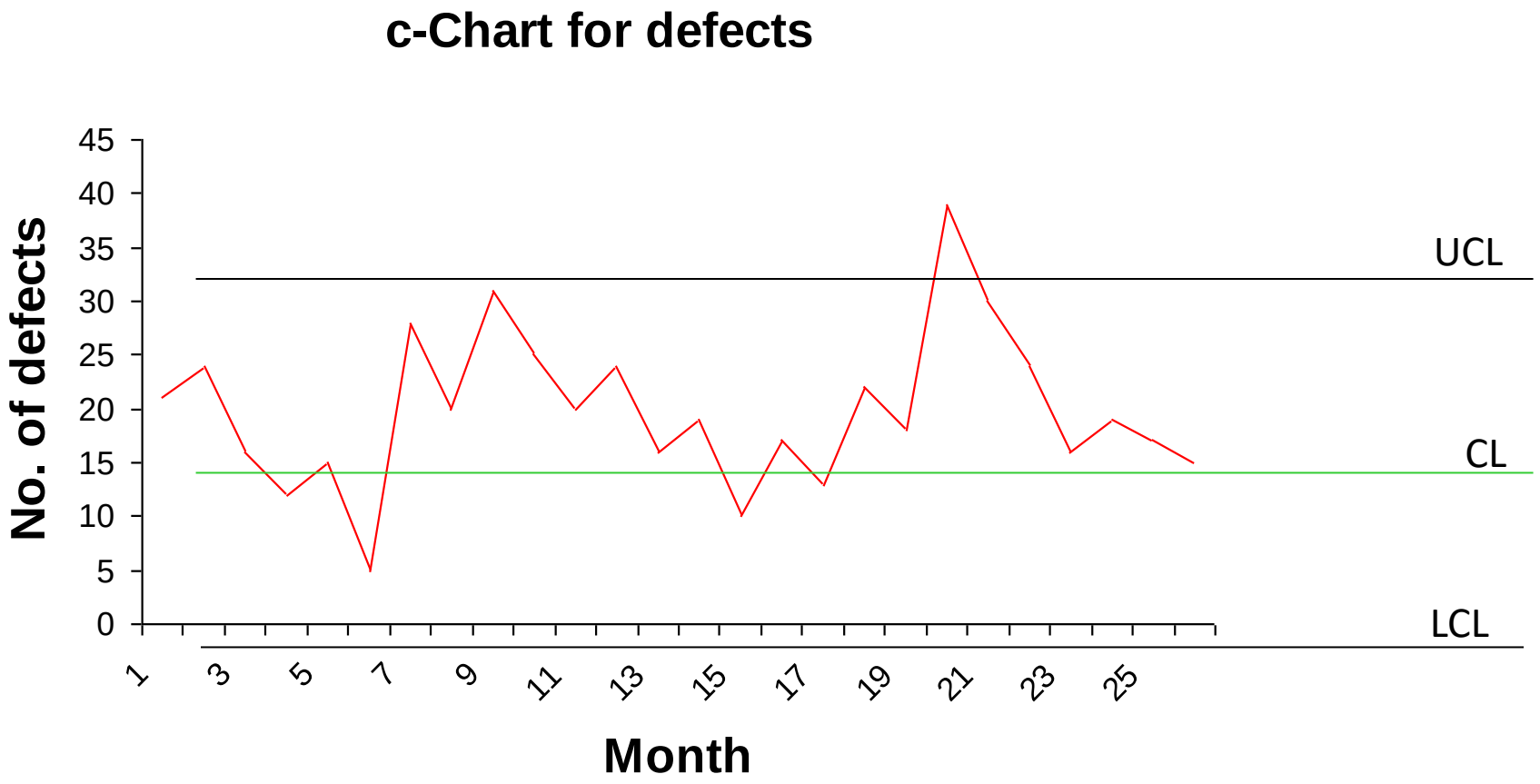
$\bar{C}$ = AVG. NO. OF DEFECTS

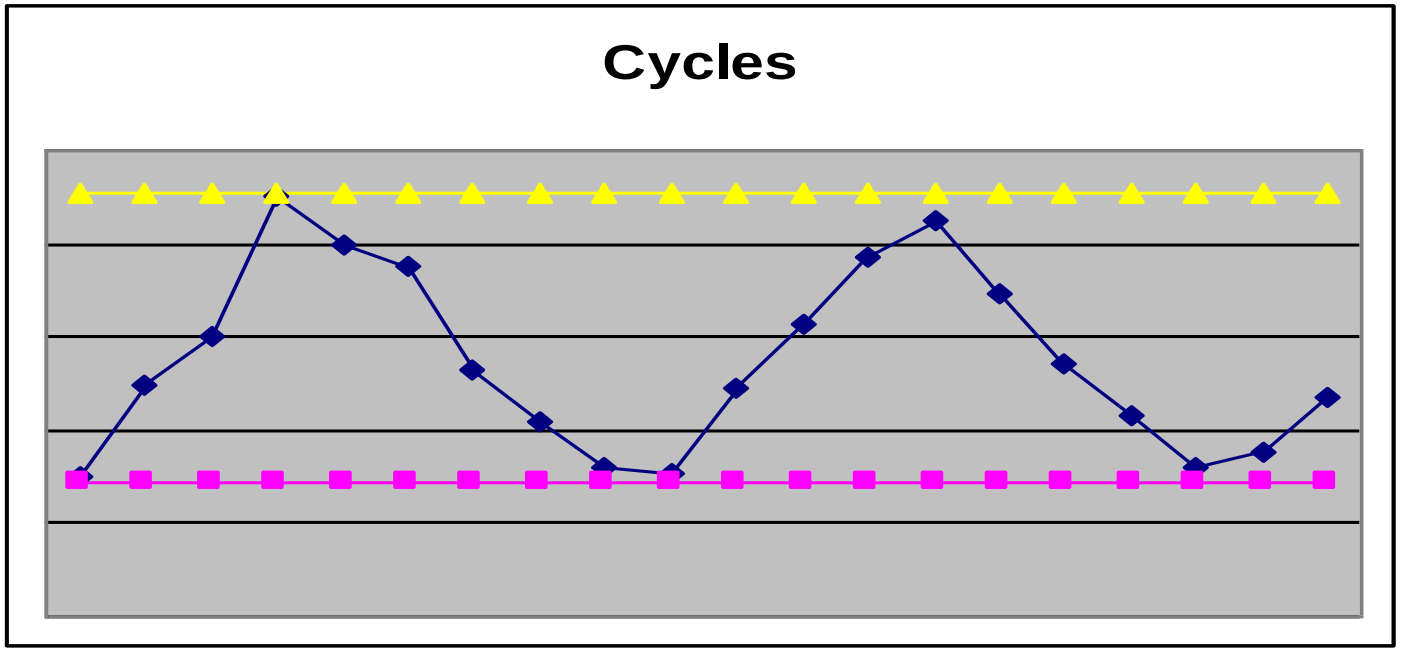
= 516 / 26 = 19.85

$UCL = \bar{C} + 3\sqrt{\bar{C}} = 19.85 + 3\sqrt{19.85} = 33.22$

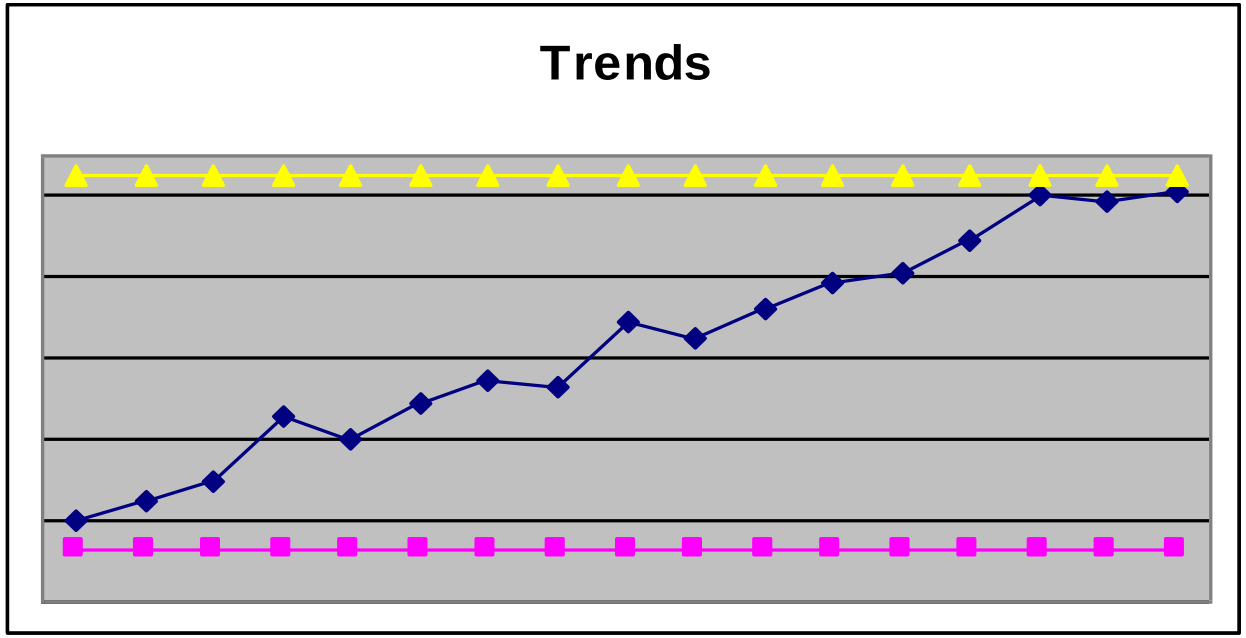
$LCL = \bar{C} - 3\sqrt{\bar{C}} = 19.85 - 3\sqrt{19.85} = 6.48$

$CL = \bar{C} = 19.85$

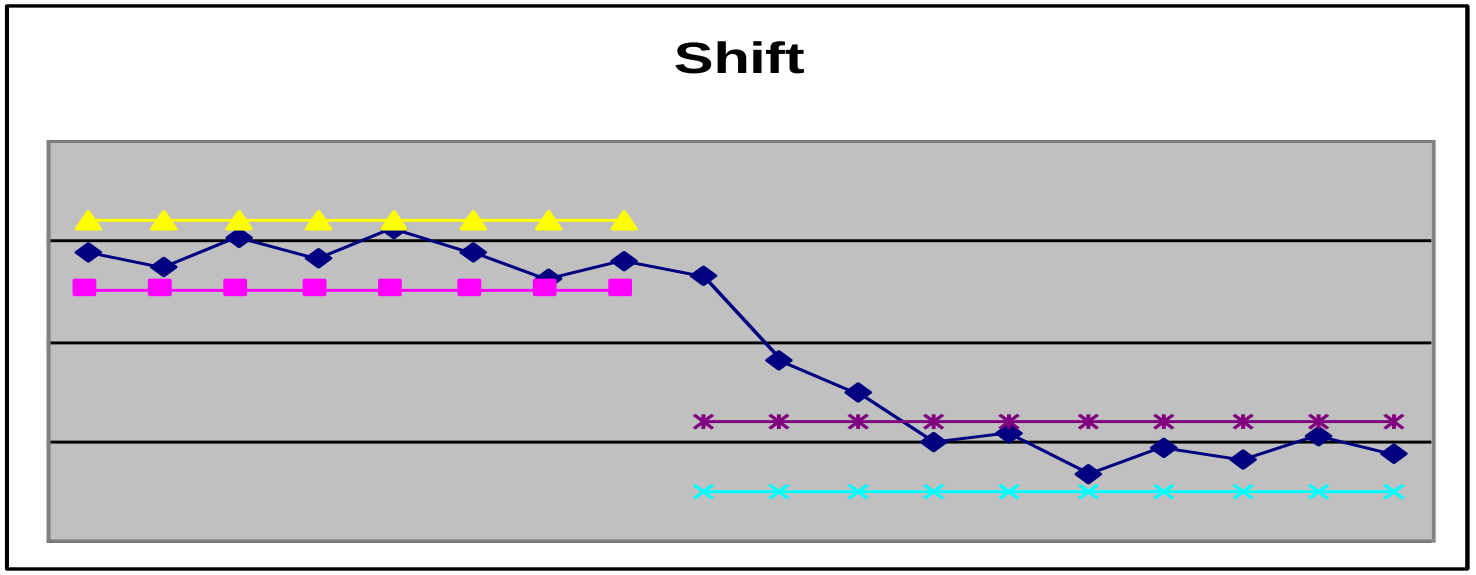

c-Chart for defects
No. of defects
45
40
35
30
25
20
15
10
5
0
UCL
CL
LCL
1
3
5
7
9
11
13
15
17
19
21
23
25
Month



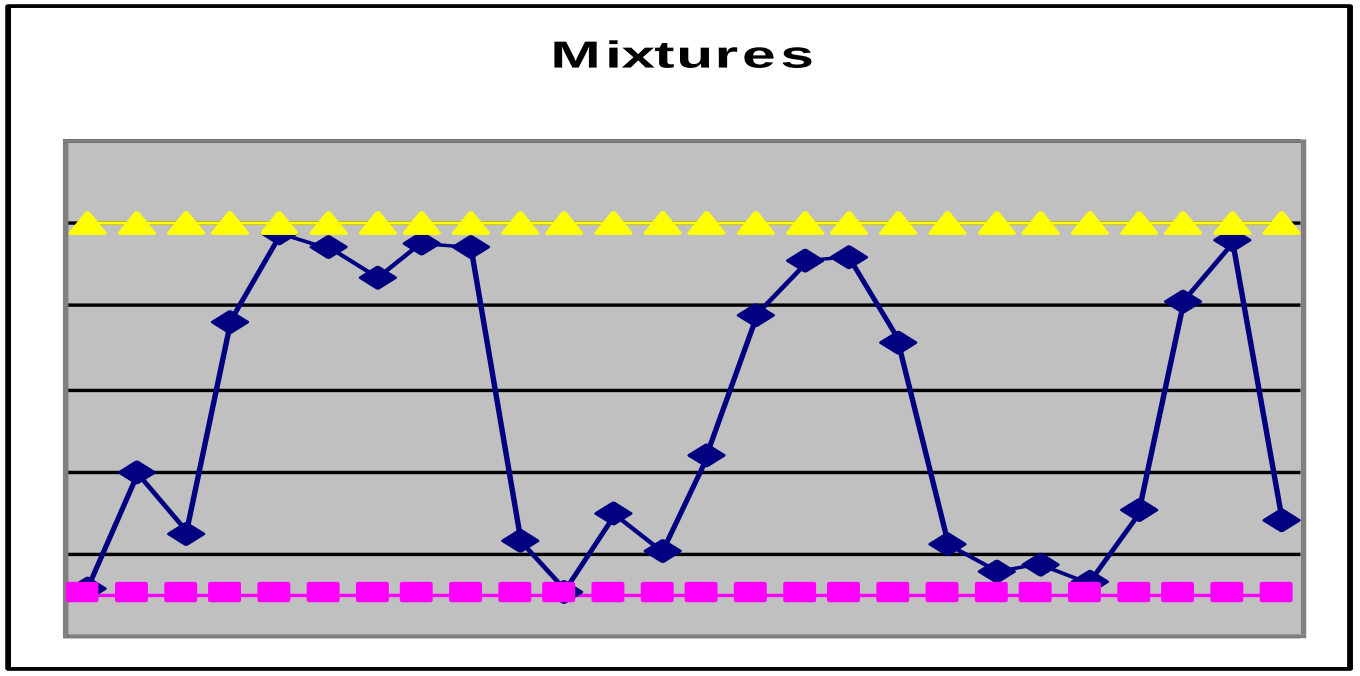
Differences between Projects/Procedures
Data Sampling Practices
Seasonal Effects
Review/Measurement Differences



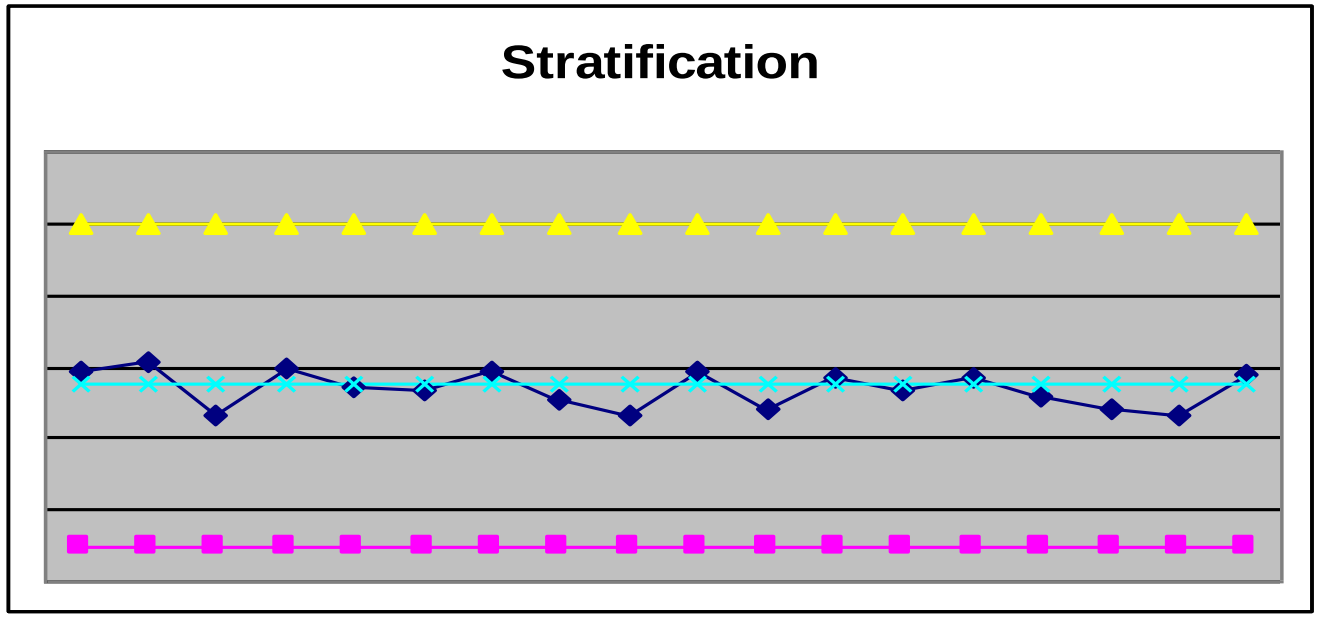
Change in Skill (Increased/Reduced) of people
Aging / Lack of Control
Change in Standards
Effect of Process Control (changes) in other areas
Introduction of different processes/people



Change of Procedure/System/People
New Process Automation/Review Tools
Change in Calibration
Inadequate/Improved Control Method



Over-adjustment of the Process
Output from 2 or more Projects / Units mixed together for analysis
Data Sampling problems (Differences in People/Tools usage)



Non-random Sampling
Data sorting at a prior state in Process
Wrong calculation of Control Limits
Failure to recalculate CLs when process capability is improved
Misplaced decimal point (wrong data entry ?)

Action Plan Matrix

Root Cause	Counter Measure	Implementation Solution	Feasibility (f)	Effectiveness (e)	f*e	Action	RPN	Control point	Recommended Action

- Process Control Plan to be implemented for sustaining improved performance

■ Resistance

Any organizational change tends to be resisted

Reasons	**Addressed by**
Self Interest	New Measures and Rewards
Fear of the Unknown	Education & Training
Differing Perceptions	Data based approach
Suspicion	Using endorsement by influential people
Conservatism	Success Stories

Ensure that people are part of the Solution and not of the problem

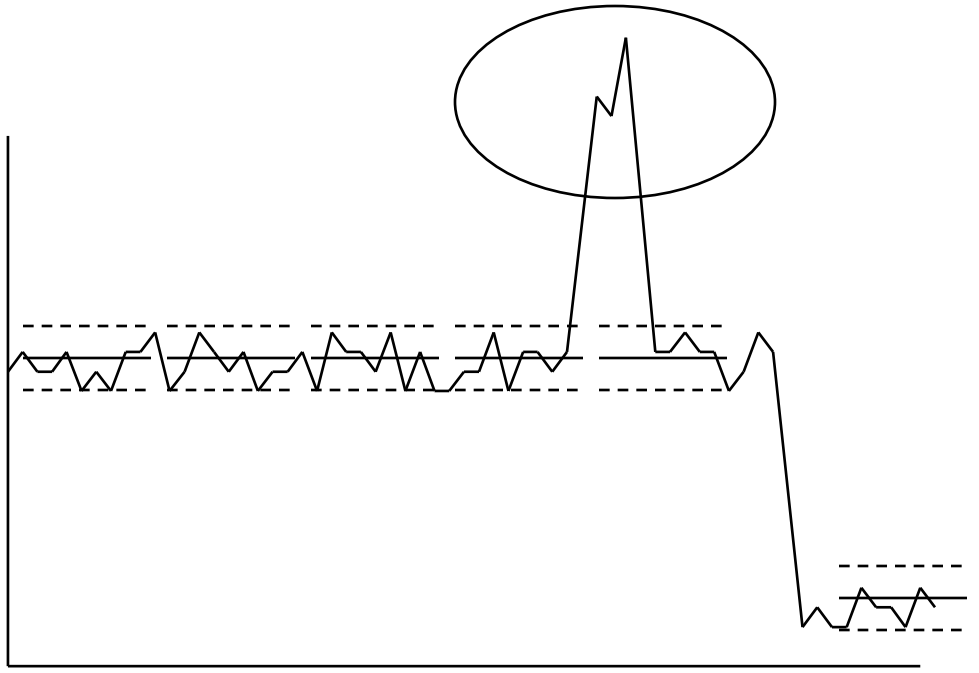
- It is important to now re-compute the process baseline in order to
  - Verify improvement levels
  - Compare new capability with predicted capability
  - Re-confirm noise levels
  - Form new confidence intervals

- It’s possible that all vital X’s are under control, but required improvement is not made
  - If X’s chosen do not functionally relate to ‘Y’
  - If some vital X’s are missed out
  - If the optimum region has not been explored completely
  - If the operating limits are not fixed properly
  - If considerable measurement error is present in both ‘X’ & ‘Y’

- If required improvement is not made, each of the above points should be explored

PLANNING → **CONTROL** → **IMPROVEMENT**

- Process / Service
  - Process Measures
  - SLAs
- People
  - Skills Requirements
  - Number / Staffing
  - Training Requirements
- Process
  - Procedures
  - Methods
- Infrastructure
  - Capabilities
  - Utilization
- Management
  - Measures
  - Policies



iSTRIVE is the process of determining what process factors should be changed, to what, and how the new process should be sustained

Questions to be answered

1. Is the Improvement (change in process performance) real and sustainable ?

2. Who will be affected?

3. What are the verified impacts /benefits ?

4. How will the solution be implemented?

5. Are the new process controls robust ?

Deliverables - Phase level

1. Financial Benefits

2. Performance Results of Improved process

3. Process control system for sustenance

4. Roll-out plan

Tools

SPC - Control Charts/Time Series	Process Control Plan
Risk Mitigation	Poka Yoke

Do's

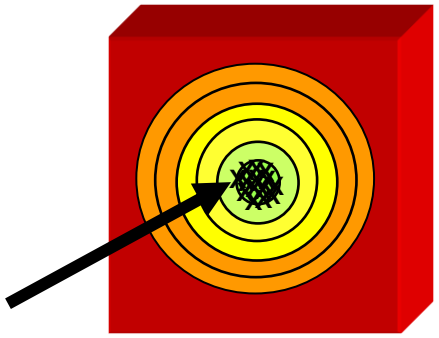
1. Incorporate Change management while planning the implementation. Again, use stakeholders buy-in for the implementation of the solution fully. Consider how the changes will benefit all the Stakeholders of the process.

2. Verify the impact of the process changes. Projections and pilots do not always translate into operational reality

Don'ts

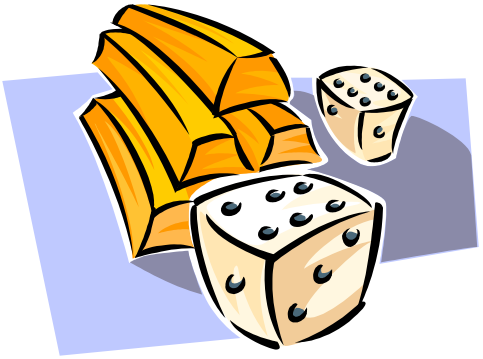
1. Assume the process changes are acceptable to all because they "make sense". Reactions are often based on emotion, concern for personal security and status. These must be identified and addressed as part of the roll-out.

2. Lose control on the Project or improvement actions just by the sight of validated results - do not give any room for complacency until the last task in the project.



Pilot Results

Confirm Benefits

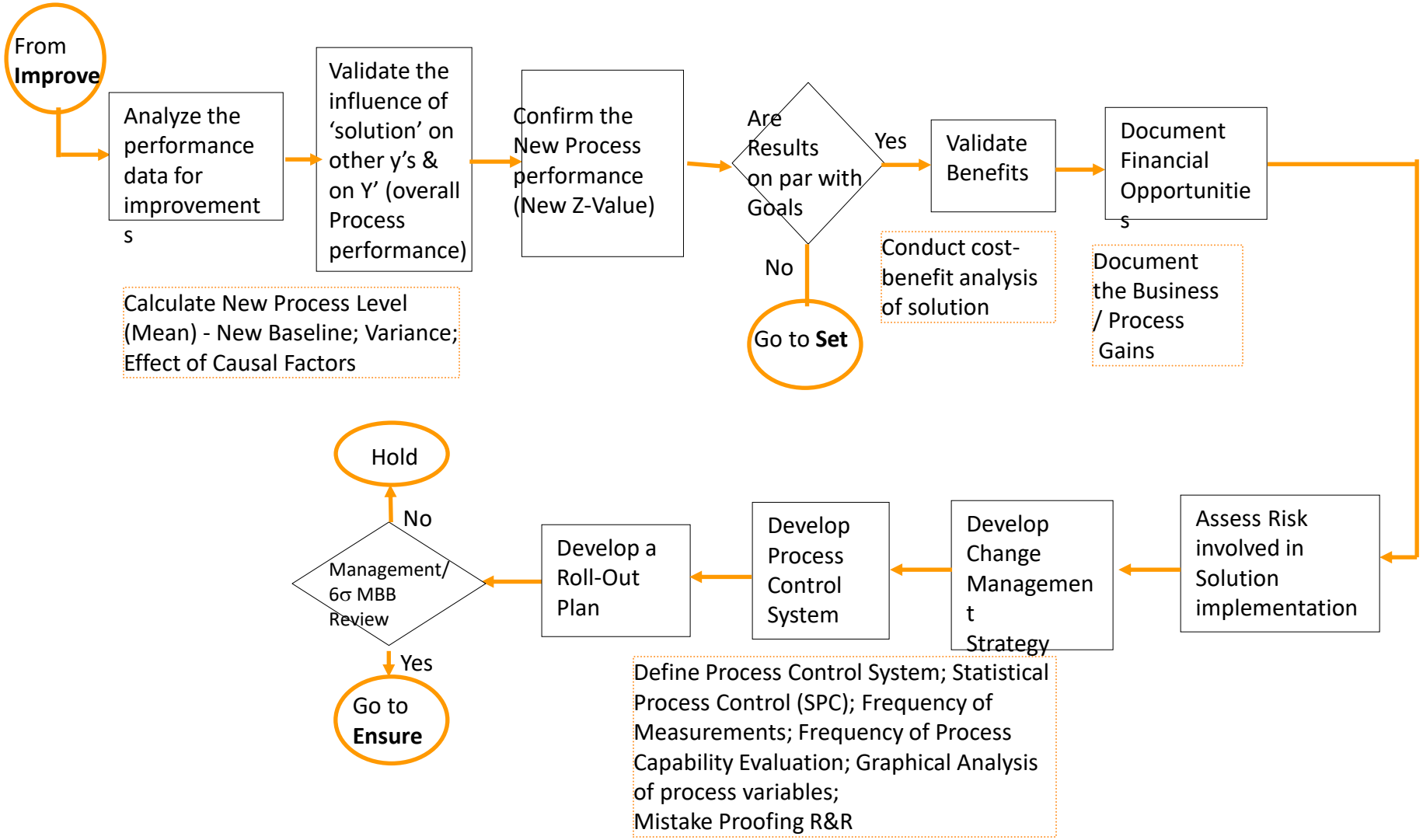


Risk Identification

Risk Mitigation Plan

Implement the Solution fully

Validate the Process Improvements


From **Improve**
Analyze the performance data for improvements
Validate the influence of 'solution' on other y's & on Y' (overall Process performance)
Calculate New Process Level (Mean) - New Baseline; Variance; Effect of Causal Factors
Confirm the New Process performance (New Z-Value)
Are Results on par with Goals
Yes
No
Go to **Set**
Validate Benefits
Conduct cost-benefit analysis of solution
Document Financial Opportunities
Document the Business / Process Gains
Assess Risk involved in Solution implementation
Develop Change Management Strategy
Develop Process Control System
Define Process Control System; Statistical Process Control (SPC); Frequency of Measurements; Frequency of Process Capability Evaluation; Graphical Analysis of process variables;
Mistake Proofing R&R
Develop a Roll-Out Plan
Management/ 6σ MBB Review
No
Yes
Hold
Go to **Ensure**

Sub-group Size	Factor for estimating process standard deviation	Factor for $\bar{X}$-R Chart	Factors for R chart		Factor for $\bar{X}$-S Chart	Factors for S chart	
n	d_2	A_2	D_3 (LCL)	D_4 (UCL)	A_3	B_3 (LCL)	B_4 (UCL)
1	1.128	2.66	0	3.27			
2	1.128	1.88	0	3.27	2.66	0	3.267
3	1.693	1.02	0	2.57	1.95	0	2.568
4	2.059	0.73	0	2.23	1.63	0	2.266
5	2.326	0.58	0	2.11	1.43	0	2.089
6	2.534	0.48	0	2.00	1.29	0.030	1.970
7	2.704	0.42	0.08	1.92	1.18	0.118	1.882
8	2.847	0.37	0.14	1.85	1.10	0.185	1.815
9	2.97	0.34	0.18	1.82	1.03	0.239	1.761
10	3.078	0.31	0.22	1.78	0.98	0.284	1.716

$$A_2 = 3 / (d_2 * n^{0.5})$$

Microsoft Word Document

(Monitoring the Invoicing Process)

Microsoft Word Document

(A Mechanism for Process Control)

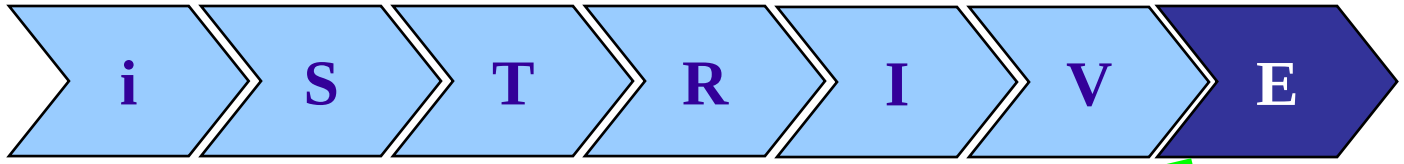
See Appendix for Case Studies

6σ

Green Belt Program

ENSURE

Module 8



Integrate Systems

- Consolidate all the changes in process from the improvement project(s)
- Revise and Update Process Documentation in Qualify, Quantify and Web-Auditor, etc.
- Document the Lessons leant and the best practices

Training & Communication

- Identify Associates to be Trained
- Conduct training on all Process Changes
- Monitor Process at Process Level (y) and Circle Level (Y)
- Summarize changes and communicate results to stakeholders.

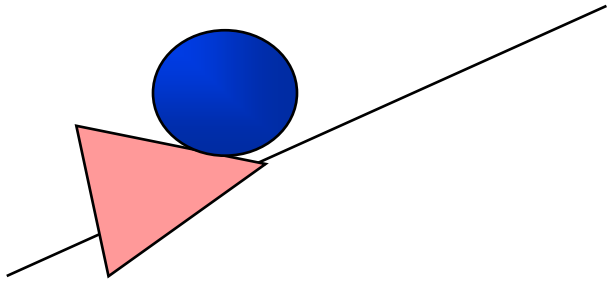
Management Approval

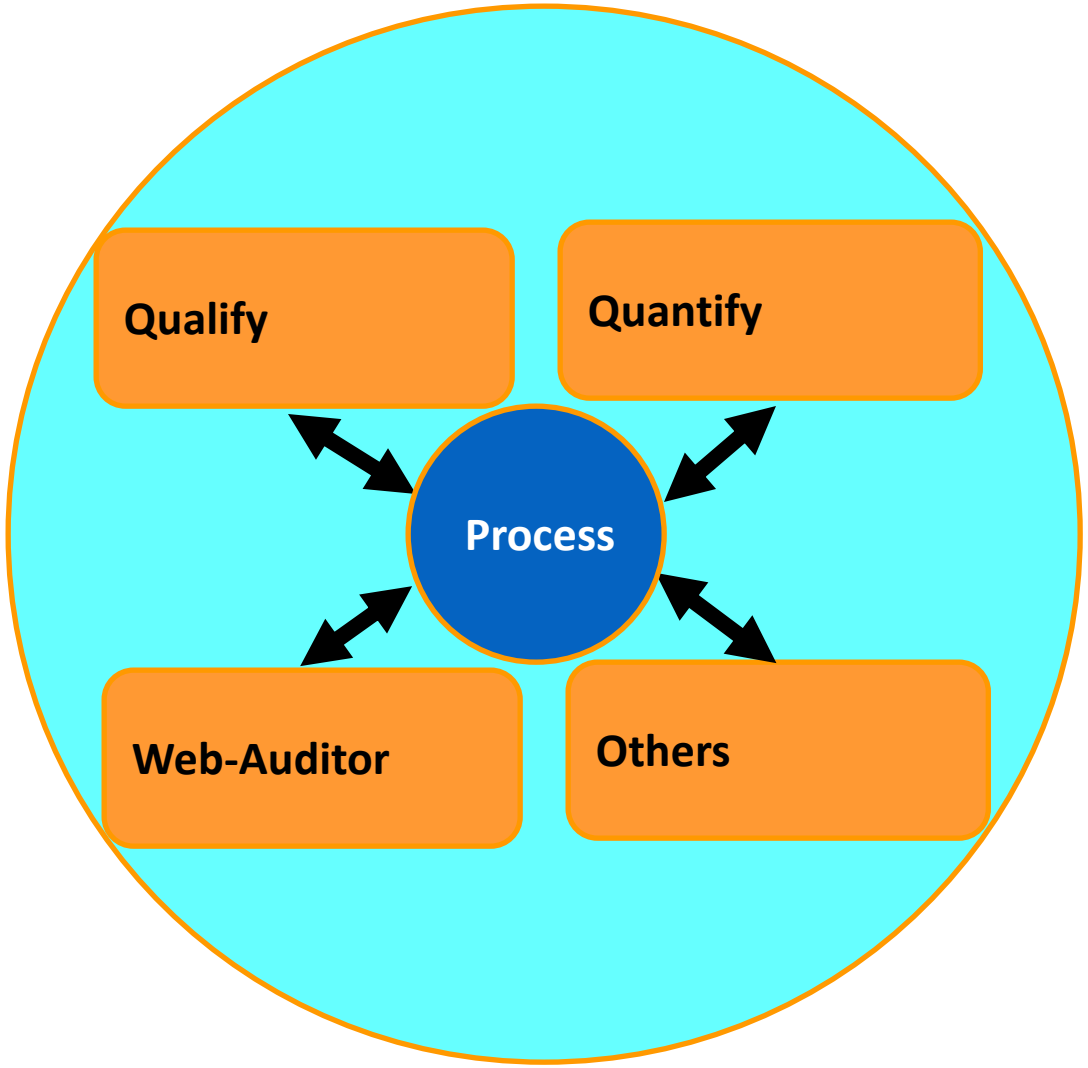
- Institutionalize Continuous improvement process by setting new benchmarks
- Review translation opportunities
- Recommend the Process for ‘Level 5’ Certification

Ensure improved process & circle performance

Business Problem → Statistical Problem → Statistical Solution → Business Solution

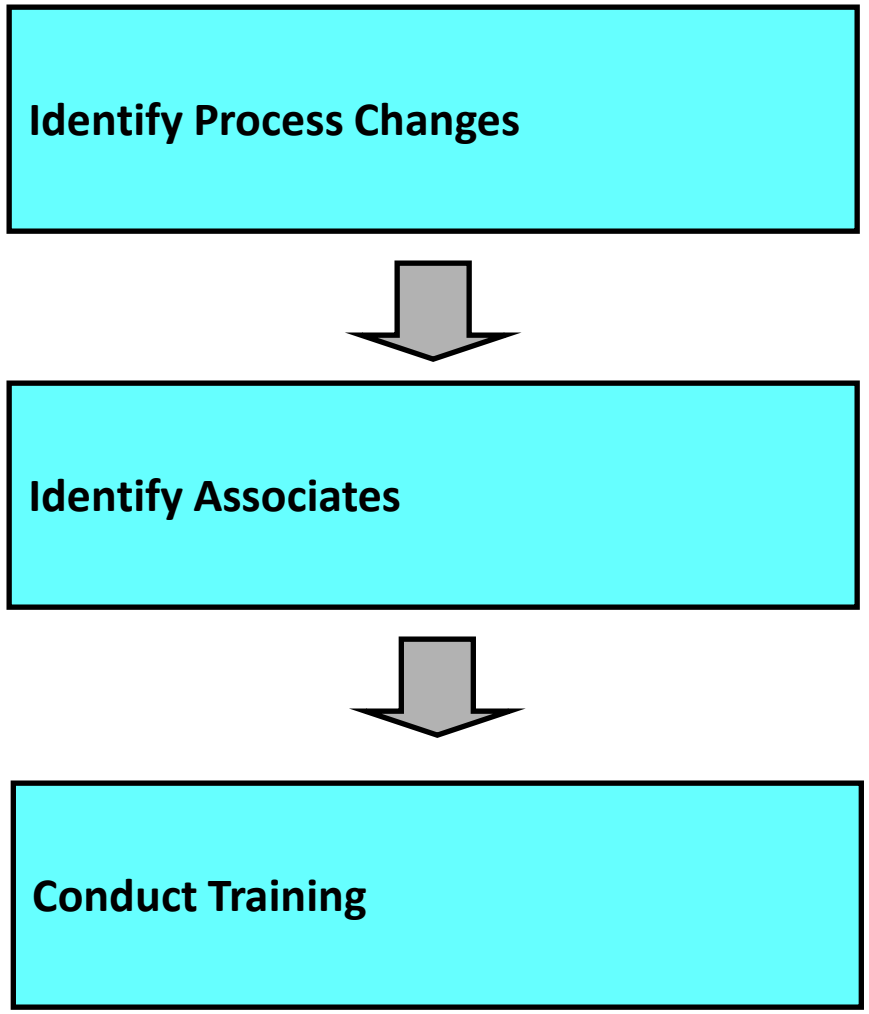
Ensure that the improvements made are sustained at Process Level (Circle Measure performance – Y)



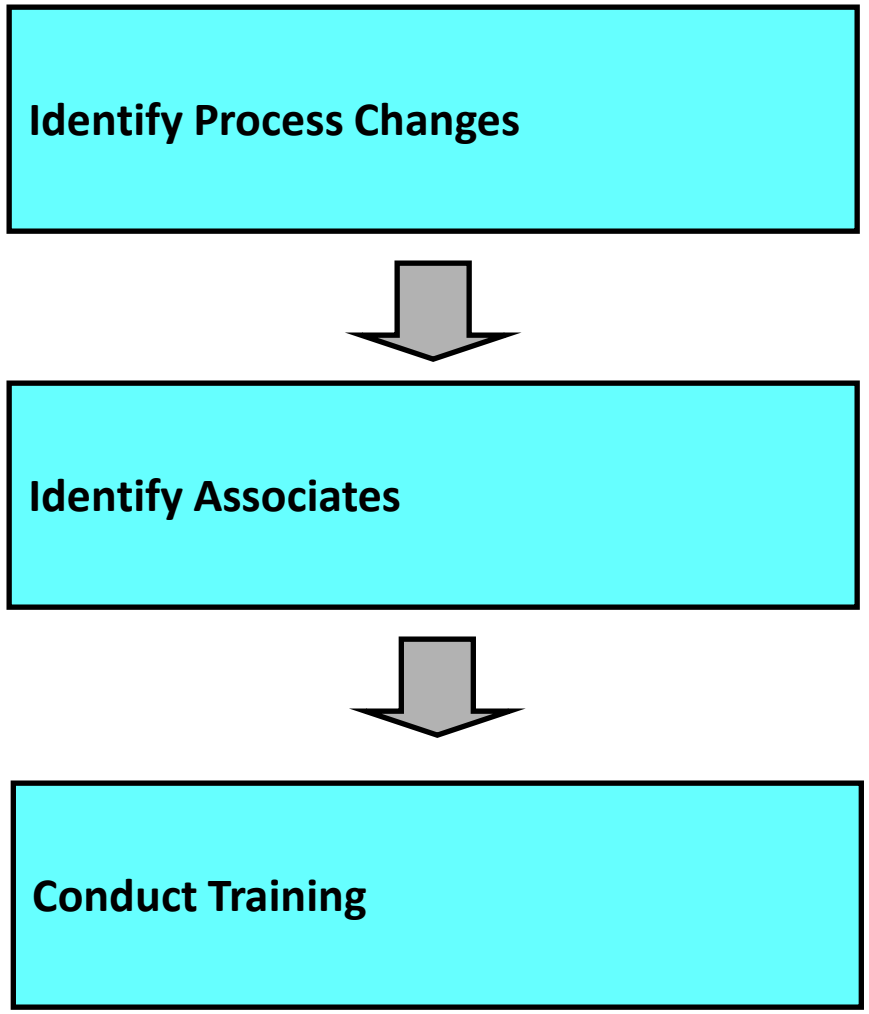

Qualify
Quantify
Process
Web-Auditor
Others

- You are ready to begin implementing what you've learned.
- Document the Lessons learnt during the course of this project
- Share the learnings with all concerned
- You need to ensure that all relevant staff in the organization is aware of and can make use of what you've learned.
- Your report and your presentations may in fact be one of the most important activities in your project.

Rubber meets the road

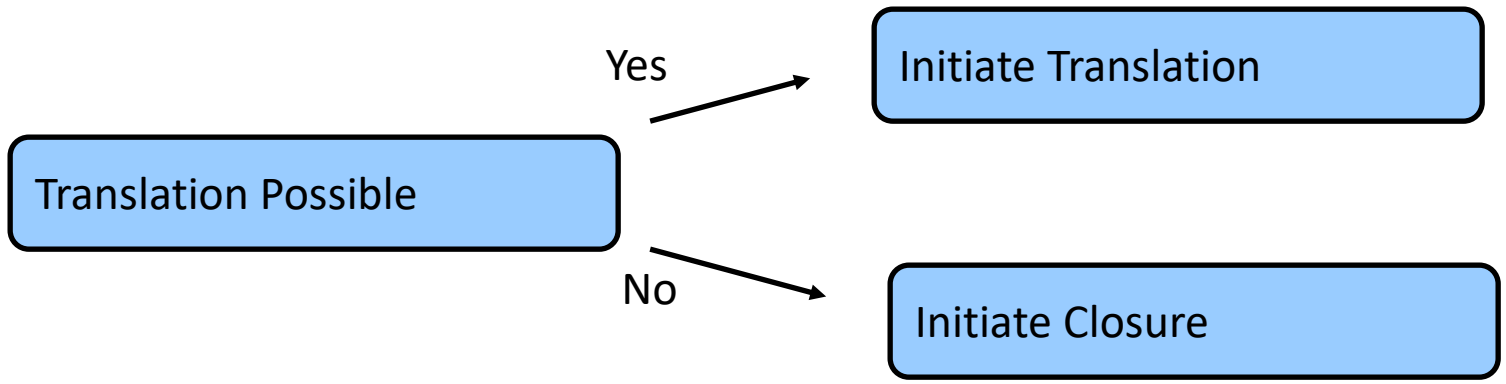

Identify Process Changes
Identify Associates
Conduct Training

- **Implement Change Management Strategy**
- **Implement the Business Solution**
- **Define Measurement (Process Measures) Aspects**


Identify Process Changes
Identify Associates
Conduct Training

- **Review Measurement (Process Measures) Aspects**
- **Monitor Process Capability**
- **Document Project Progress & Results**
- **Document Lessons Learnt**

- **Look for Similar Processes**
- **Evaluate Translation Possibilities**



- **File for Level 5 Certification**
- **Set New Benchmark**
- **Continue on Improvement Path**
- **Formally Close Project**

Institutionalize Continuous Improvement

Questions to be answered

1. How will the new process changes be integrated at the Process/Circle level?

2. Are the changes documented ?

3. Is there adequate training to concerned associates on the process changes?

4. How will the change be monitored, sustained & institutionalized ?

5. Are the process changes communicated?

Deliverables - Phase level

1. Process Documentation & Integration

2. Training of Identified Associates

3. Communication to AICS

4. Translation Opportunities Documented

5. Project Closure with Management Approval and Recognition.

Tools

Bench-marking **Qualify/QMS**
Quantify **Web Auditor**

Do's

1. Document and Communicate the process changes to the stakeholders.

2. Integrate the changes at the process level in Qualify, Quantify,Web-auditor, etc.

3. Identify the Do's & Don'ts and impart training

4. Evaluate the improvement for Translation Opportunities

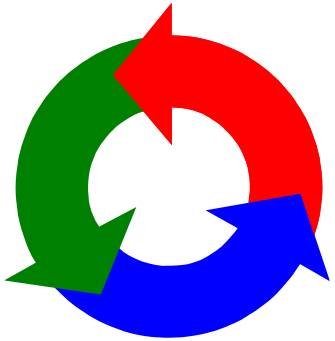
Don'ts

1.Recommend project closure until the project documentation is completed with the best practices /lessons learnt

2. Recommend project closure until Circle PI and Process PI reflect the Improvements achieved in Project y's

3. Recommend project closure until the verification of "translation opportunities" of the present success is done

Update Documentation



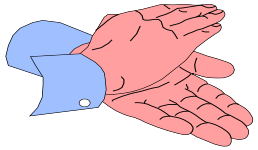
Integration at Process-level (Qualify, Quantify, Web Auditor etc)

Conduct Training

Measures & Targets and New Baseline (Monitor results to Sustain Benefits)

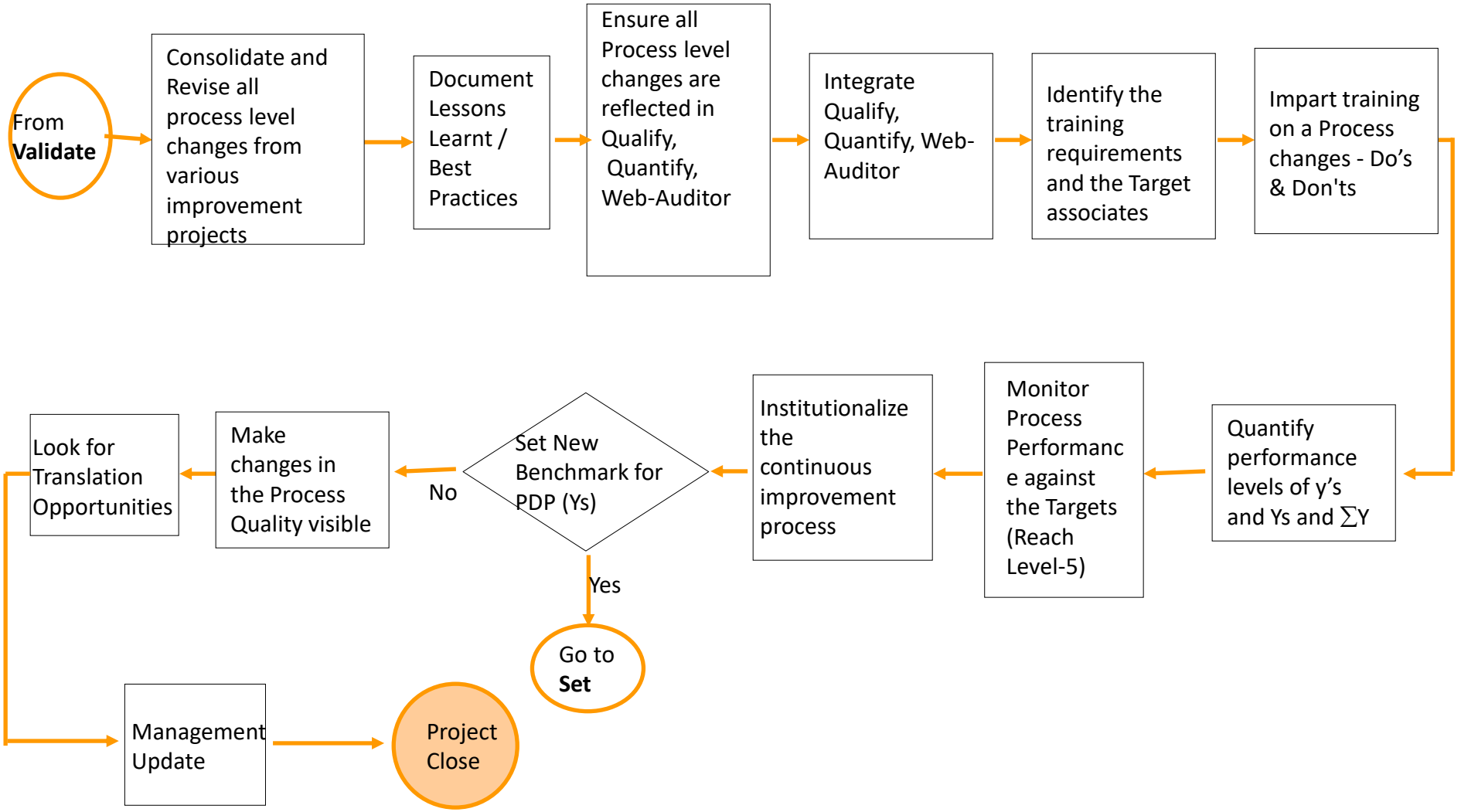
Translation Opportunities

Communicate Results



Celebration

Ensure improved process & circle performance


From **Validate**
Consolidate and Revise all process level changes from various improvement projects
Document Lessons Learnt / Best Practices
Ensure all Process level changes are reflected in Qualify, Quantify, Web-Auditor
Integrate Qualify, Quantify, Web-Auditor
Identify the training requirements and the Target associates
Impart training on a Process changes - Do's & Don'ts
Quantify performance levels of y's and Ys and ∑Y
Monitor Process Performance against the Targets (Reach Level-5)
Institutionalize the continuous improvement process
Set New Benchmark for PDP (Ys)
No
Yes
Make changes in the Process Quality visible
Look for Translation Opportunities
Go to **Set**
Management Update
Project Close

This is not the End…

Neither is this the beginning of an End…

Perhaps, it is the end of a Beginning...