### **Continuous Improvement Toolkit**

### **Scatter Plots**



### The Continuous Improvement Map

Managing	Decidir	Deciding & Selecting		Planning & Project Management*		
Risk PDPC	Decision Balance Shee	t Importance-Urgency	Mapping D	aily Planning	PERT/CPM	
FMEA RAID Log*	Force Field Analysis	Cost Benefit Analy <mark>si</mark>	s <u>MOST</u>	RACI Matrix	Activity Networks	
Risk Assessment*	Break-even Analysis	Voting TPN Analysis	SWOT	Analysis Sta	keholder Analysis	
Fault Tree Analysis	Decision Tree Pick Char	t Four Field Matri <mark>x</mark>	Project Ch	arter Impro	vement Roadmaps	
Traffic Light Assessmer	t Critical-to Tree QFD	Portfolio Matrix	PDC	A Policy Deploy	ment Gantt Charts	
Lean Measures Ka	ano Analysis Matrix Diagran	n Paired Comparison	DMAIC Kaiz	en Events Co	entrol Planning	
Bottleneck Analysis**	Cost of Quality* Pugh Matrix	Prioritization Matrix	A3 Thinking	Standard work	Document control	
O Process Vield	EE <u>KPIs</u> Pareto Anal	ysis C&E Matrix	rstanding	Cross Training	Implementing	
D Conchility Indiana	escriptive Statistics ANOVA	Chi-Square Caus	e & Effect	Value Analysis	Solutions**	
	Probability Distributions Hyp	oothesis Testing Design	n of Experiment	Mistake Proof	ing Ergonomics	
Gap Analysis <sup>*</sup> His	tograms & Boxplots Multi v	ari Studies Confidence	ce Intervals Si	mulation TPN	Automation	
Reliability Analysis	Fraphical Analysis Scatter P	lots Correlation R	Regression	Pull Flow	Just in Time	
Performance MS	A Run Charts 5 Whys	Root Cause Analysis	Data Snooping	Visual Manage	ment 5S	
Benchmarking** (	Control Charts Fishbon	e Diagram Tree Diagrar	n* SIPOC*	Waste Analysis	Quick Changeover	
Data collection planner*	Sampling Morphological	Analysis How-How Dia	agram** Prod	<mark>cess Re</mark> design	Time Value Map	
Check Sheets Intervie	ws Brainstorming SCAM	IPER** Attribute Analy	sis Spaghett	i Diagram Valu	ue Stream Mapping	
Questionnaires Focu	Affinity Diagram	Relationship Mappir	ng* Flow Pr	ocess Charts	Service Blueprints	
Data	Mind Mapping*	Lateral Thinking	Flowcharting	IDEF0 P	rocess Mapping	
Collection Obse	Suggestion systems	S Creating Ideas	Desig	ning & Analyz	ing Processes	

- □ A diagram that shows whether two variables are correlated.
- Shows patterns in the relationship that cannot be seen by just looking at the data.
- Used as a first step in analyzing correlation between pairs of variables before conducting advanced statistical analyses.
- Works with both continuous and count data.



### Example:

- A line manager for example may want to check the relationship between:
  - The number of training hours and employee productivity.
  - The number of defects and the experience of the staff.
  - The equipment downtime and its cost of maintenance.





### **Other Examples:**

- **The relationship between:** 
  - Driving speed and fuel consumption.
  - The number of people working on a shift and the average answer time in a call center.
  - The number of years of education someone has and the annual income of that person.







### Continuous Improvement Toolkit . www.citoolkit.com

## - Scatter Plots

- Used to visually investigate the relationship between two variables.
- Used to verify that a change in one variable can affect the other variable.
- Helps detecting the primary factors that are really causing a problem.
- Helps eliminating non-critical factors from consideration.
- Used to determine the strength of the relationship.
- Used with statistical tools to support or reject hypotheses about the data.



□ When comparing an input with an output variable.

- The **explanatory variable** is normally placed on the horizontal axis.
- The **response variable** is placed on the vertical axis.



You may also compare two input or output variables

### How to Construct a Scatter Plot?

- □ Collect the two paired sets of data.
- Create a summary table of the data.
- Draw and label the horizontal and vertical axes.
- Plot the data pairs on the diagram by placing a dot at the intersection of each data pair.
- □ Look at how the two variables vary together.



Scatter plots can indicate several types of correlation:

No correlation when the data points are<br/>scattered randomly without showing<br/>any particular pattern $\times \times \times \times \times \times$ <br/> $\times \times \times \times \times$ 



Scatter plots can indicate several types of correlation:

A **positive correlation** occurs when the values of one variable increase as the values of the other also increase



The fitted line slopes from bottom left to top right

Scatter plots can indicate several types of correlation:

A **negative correlation** occurs when the values of one variable increase as the values



The fitted line slopes from upper left to lower right

Scatter plots can indicate several types of correlation:

# Scatter plots can also indicate **nonlinear relationships** between variables $\begin{array}{c} \times & \times \\ \times & \times \\ \times & \times \\ \times & \times \end{array}$



### **Example:**



The **volume** and the **diameter** of sample trees in a forest

**Example** – An analysis that was conducted for diagnosing the presence of diabetes at a workplace.



The population is generally young (75.8% are below thirty).

This scatter plot illustrates that there is no obvious relationship between age and glucose levels.

High glucose levels are found in all ages above twenty, and normal glucose levels are found in higher ages.

### **Matrix Plots:**

- Summarizes the relationship between several variables.
- □ Produces a scatter plot for every combination of variables.
- Allows to visually assess the variables that might be related.



### **Example:**



There is a relationship between the years of experience and salaries The number of publications does not appear to be correlated with the years of experience

### **Further Information:**

- When the relationship is not so clear, Correlation can be used to help determine if a relationship exists between the variables.
   Regression techniques go a step further by defining the relationship in a mathematical format.
- Be careful before concluding that there is a direct cause-andeffect relationship between the variables. There might be a third factor that is causing the change in the two variables.
- You can also illustrate a stratification factor in the scatter plot.
   For example, the relationship between a process output and a process input for two different settings.

**Example -** The amount of sales per month generated at two locations:

- □ The plotted points form a negative slope.
- The sales at location B is inversely related to the sales at location A.
- Does this mean that location A caused the decrease in sales at location B, or vice versa?

**Answer:** Not necessarily, unless the two locations are direct competitors.



### Where Does It Fit?

