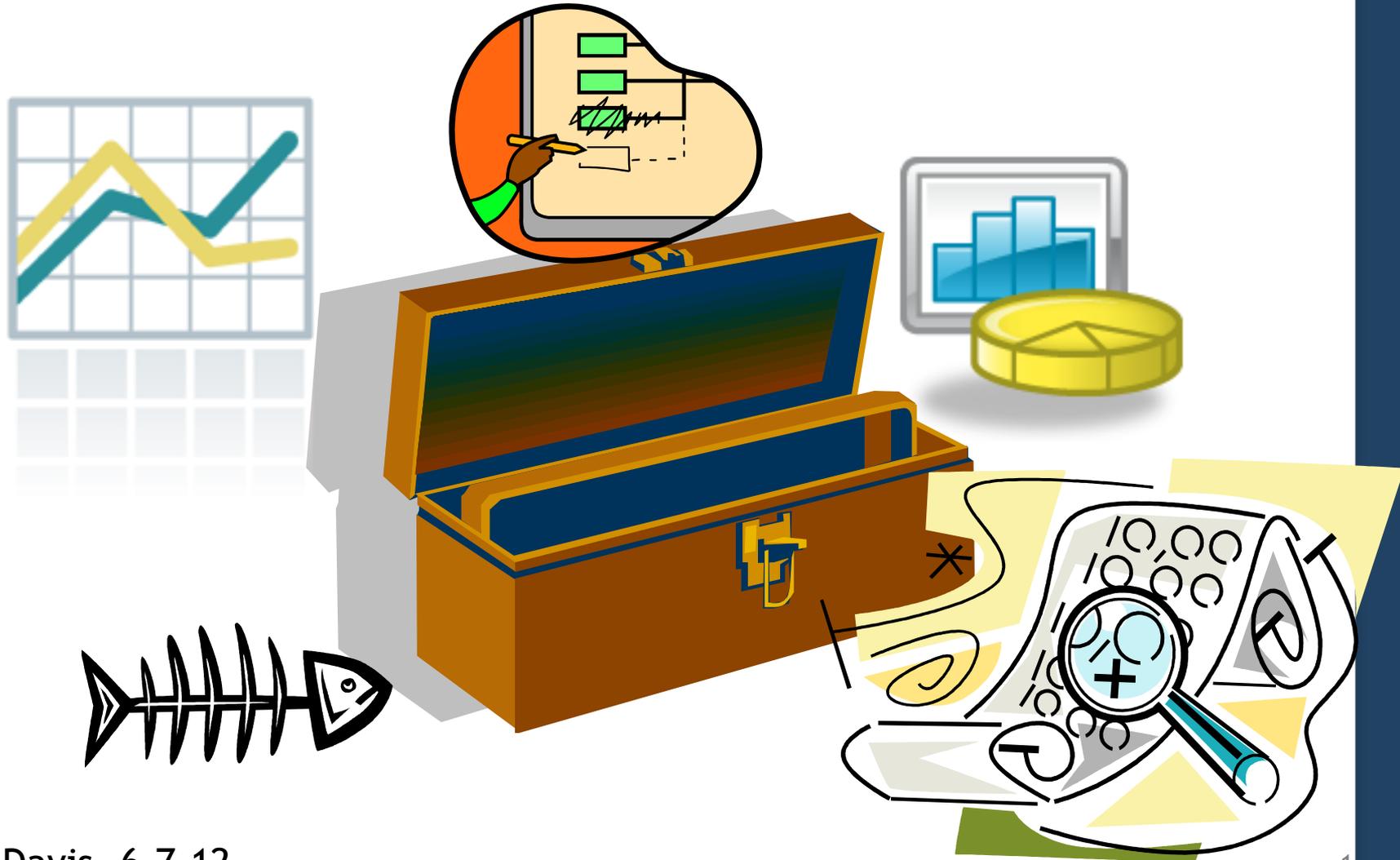


# QI Analysis Tool Box



Scott Davis, 6-7-12

# Analysis Tools

## Summary

- **Flow Chart:** Helps document current process flow and I.D. potential gaps, flaws, delays, uncertainties
- **Cause & Effect Diagram:** Helps conceptualize and sort possible contributors/causes
- **Pareto Diagram:** Data tool used to determine major contributors to a “defect” problem (pie charts can work, too)
- **Histogram:** Data tool useful for identifying variation in “cycle time” or other continuous variables (rates, volume, etc.)
- **Line/trend chart:** Data tool useful in identifying patterns of variation over time (also your key Evaluation/Control data tool)
- **Scatter Plot:** Data tool useful for finding associations between two continuous variables
- **Box & Whisker Plot:** Data tool useful for identifying variation for continuous variables, especially when sample size is small and/or you want to compare multiple samples side by side

# QI Analysis Tools

 A Lot

## When will I use them?

 A little

	Assess	Define	Analyze (early)	Analyze (late)	Change	Evaluate
Flow Charting						
Stratification						
Pareto Diagrams						
Histograms						
Box Plots						
Scatter Plots						
Trend Charts						
C.E.D.s						

# Analysis Tool (concept) Stratification



aka ... “Cutting” the data ...

# Analysis Tool (concept) “Stratification”

- To make efficient progress on a complex problem you may break it apart
- Variation may be associated with different factors
- Sorting among those factors will help you to know where to focus and have the “biggest bang”

# Analysis Tool (concept) “Stratification”

Results might vary by ...

Time

Day of week

Month/Season

Batch

People/Groups

Customer

Vendor

ETC!

# Stratification

## Beware Pitfalls

- **Confusing association with cause**

Just because things are associated does not mean there is a causal relationship

Just because you've demonstrated an association doesn't mean you know what to fix

- **Doing too little**

Not taking the opportunity to break the problem down and therefore working on more than you need to

Jumping to conclusions

Rounding up the “usual suspects”

- **Doing too much**

Breaking the problem into so many parts you are no longer having an impact on the whole

# Analysis Tool (concept) “Stratification”

- To make efficient progress on a complex problem you may break it apart
- Variation may be associated with different factors
- Sorting among those factors will help you to know where to focus and have the “biggest bang”

# Analysis Tool (concept) “Stratification”

Results might vary by ...

Time

Day of week

Month/Season

Batch

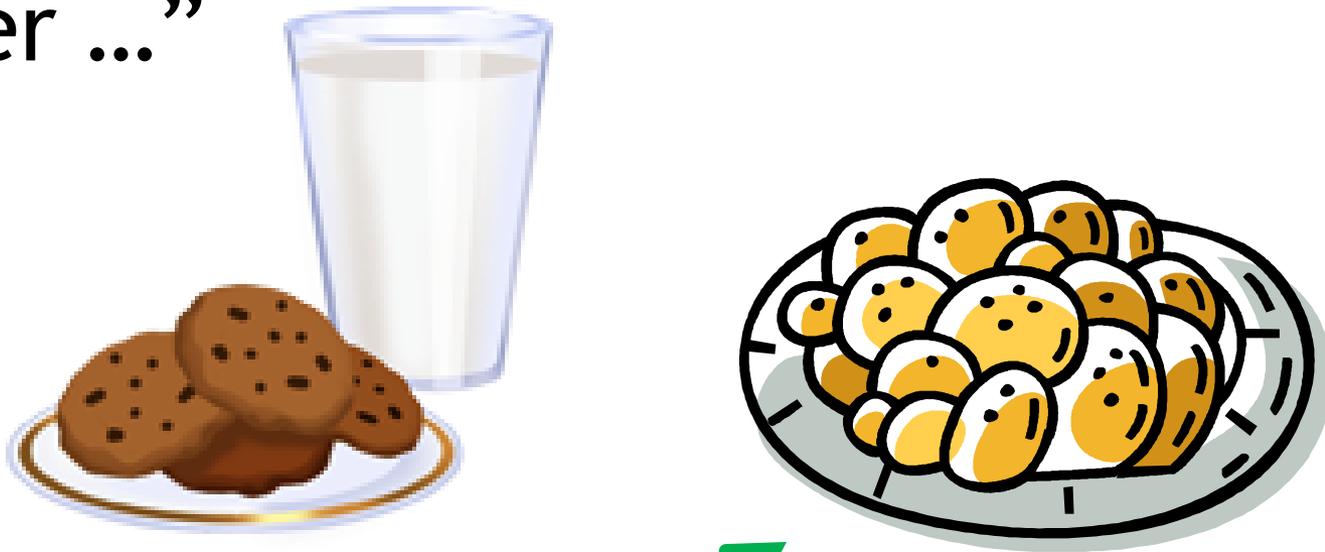
People/Groups

Customer

Vendor

ETC!

“One of these things is not like the other ...”



*These chocolate chip cookies are different shades*

**Why might they vary?**

# Stratification

## Beware Pitfalls

- **Confusing association with cause**

Just because things are associated does not mean there is a causal relationship

Just because you've demonstrated an association doesn't mean you know what to fix

- **Doing too little**

Not taking the opportunity to break the problem down and therefore working on more than you need to

Jumping to conclusions

Rounding up the “usual suspects”

- **Doing too much**

Breaking the problem into so many parts you are no longer having an impact on the whole

# Data Collection



# Analysis: Data Collection Planning

Information is an

*answer to a thoughtful question*

- Not all data is information
- Collecting data can be time consuming and expensive
- Increase success and reduce frustration through careful planning and thinking ahead

# Data Collection Planning Matrix

Question to be answered	Possible source of information	Method/tool for gathering info	Sample size required	How info will be displayed

# Data Collection Tool

## Check Sheet Errors

	Y						
	X						
	X						
	X				X		
	Y				X		
Y	X		Y		X		
X	X		Y		X		
X	X	X	X	Y	X		
<b>J.J</b> .	<b>K.C</b> .	<b>A.A</b> .	<b>L.H</b> .	<b>S.D.</b>	<b>D.H</b> .	<b>T.M.</b>	

**Teams**

X = error cost < \$100  
Y = error cost > \$101

# Data Collection Tool

## Data Sheet

Date \_\_\_\_\_

Employee \_\_\_\_\_

Order # \_\_\_\_\_

Team \_\_\_\_\_

Order type \_\_\_\_\_

Vendor \_\_\_\_\_

Customer \_\_\_\_\_

Error description:

Cost: \_\_\_\_\_

# “Organic” Data Collection

Look for “natural” collection points

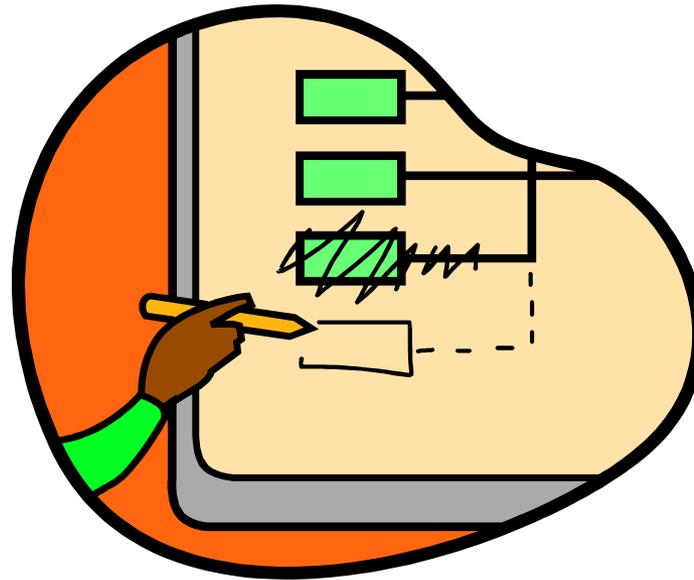
- Inspections/reviews
- data input points
- Forms we already complete
- “Artifacts” ... we collect A LOT of information that we may not use or report on

# Data Collection

## Beware Pitfalls

- **Test your data collection tool!**
- When in doubt - collect more
  - so you don't have to go back later
- Beware omission bias
  - Try to get representative sample
  - Your data collection needs to be consistent
  - Your confidence in the data will increase with the sample size
- Beware “paralysis by analysis”
  - Data is rarely perfect
  - QI rarely needs as much confidence/certainty as hard science
- Minimize effort
- Remember ... risk, cost, complexity

# Process Flow Charting



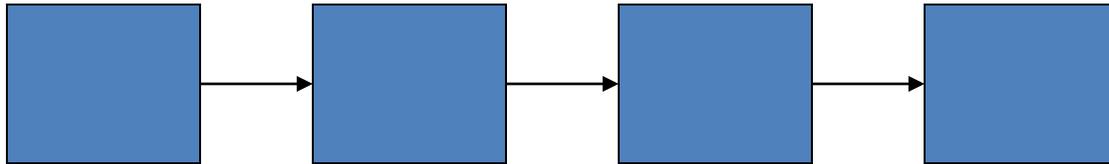
# Flow Charting

## When would you use this tool?

- Document existing process
- Establishing a standard
- Analyzing existing process(es)
- Developing solutions
- Communicating solutions/new process

# Analysis Tool

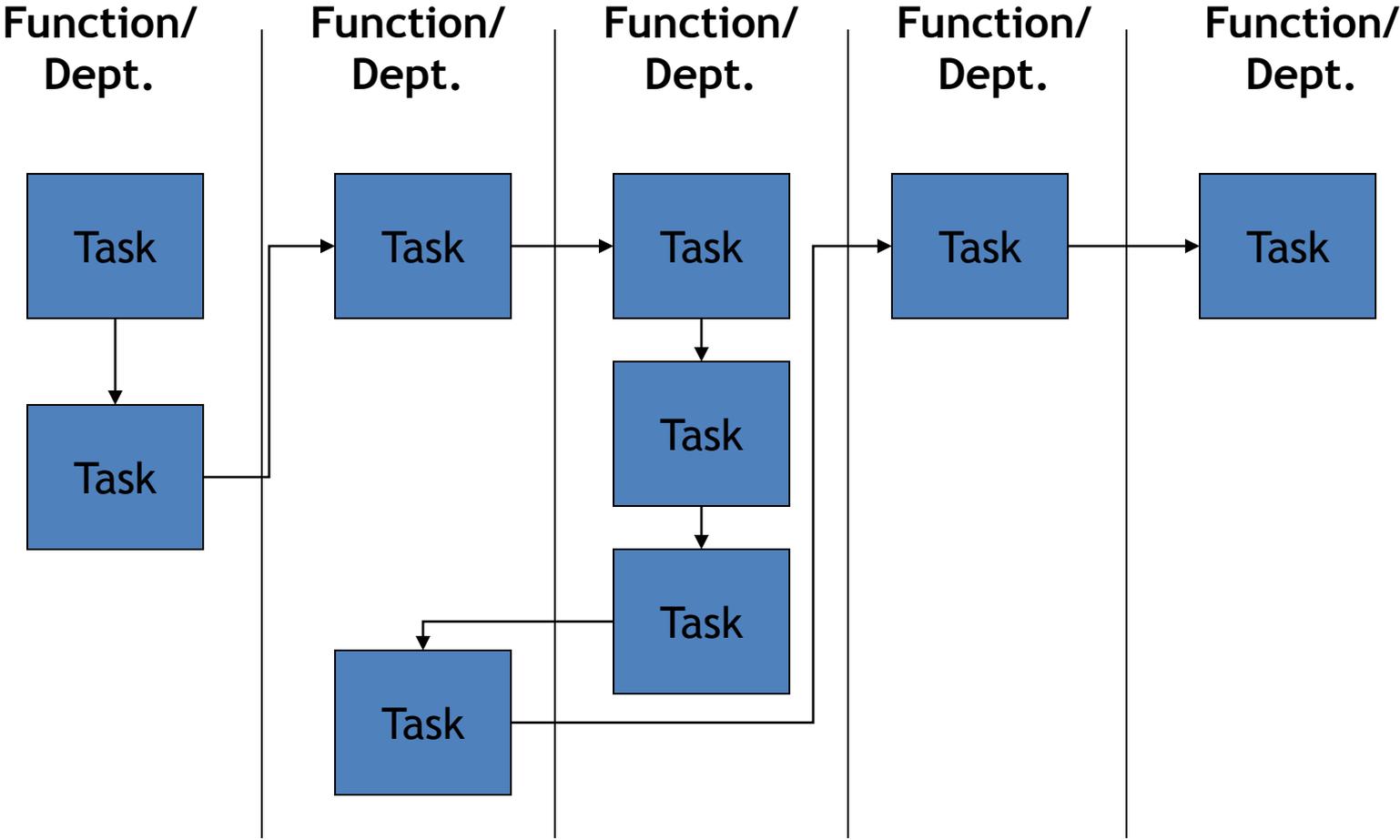
## High Level Flow Chart



- Documents start and end of process
- Documents primary objectives of process
- 4-7 steps (typically)
- More conceptual

# Analysis Tool

## Matrix Flow



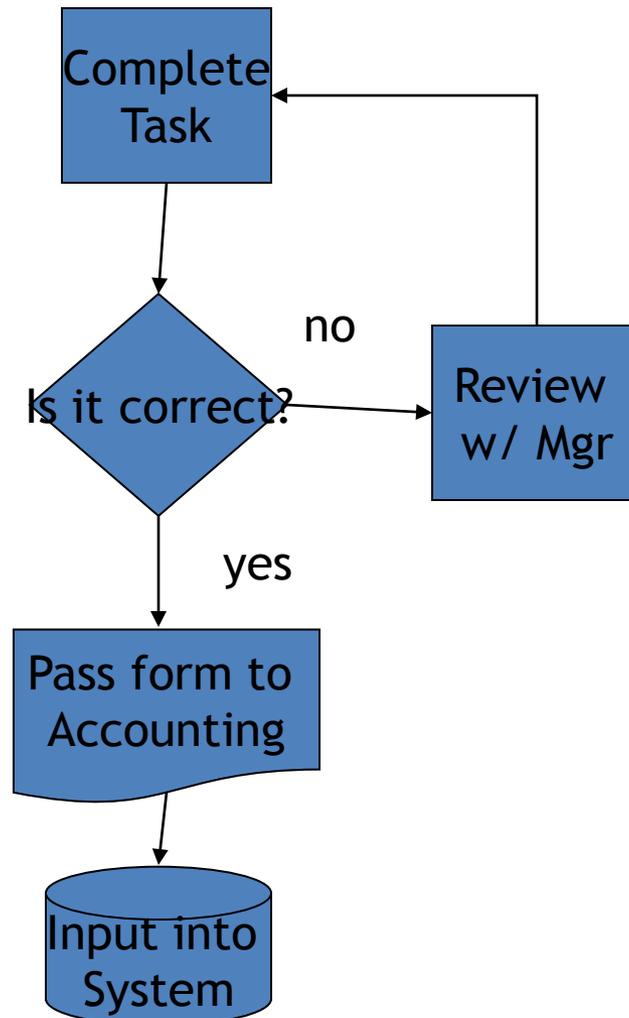
# Analysis Tool

## Matrix Flow

- Documents objectives of process and what department/function is accountable
- Documents the major steps necessary to complete the objectives
- Documents relationships and hand-offs between functions/departments
- Still somewhat conceptual
- Flexible ... can add detail as needed
- Easy for others to follow / understand

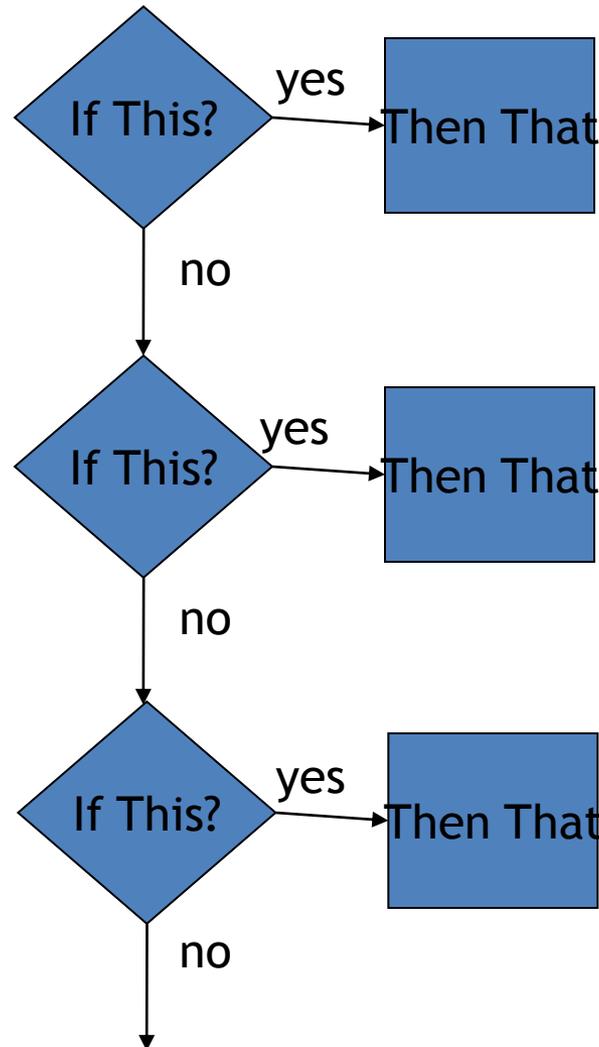
# Analysis Tool

## Detailed Process Flow Chart



# Analysis Tool

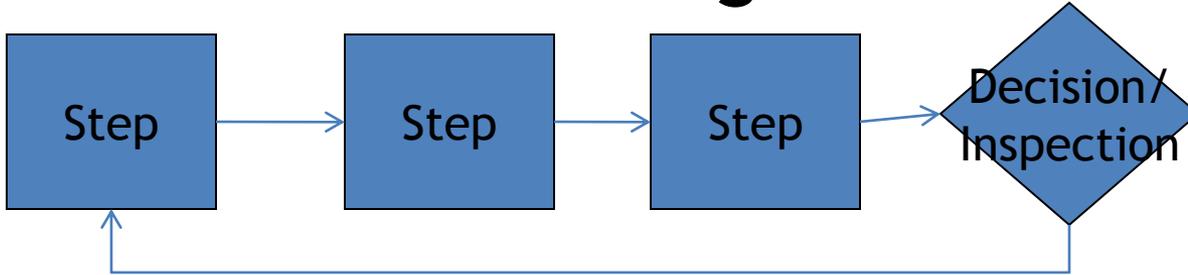
## Detailed Process Flow Chart - “Algorithm” Format



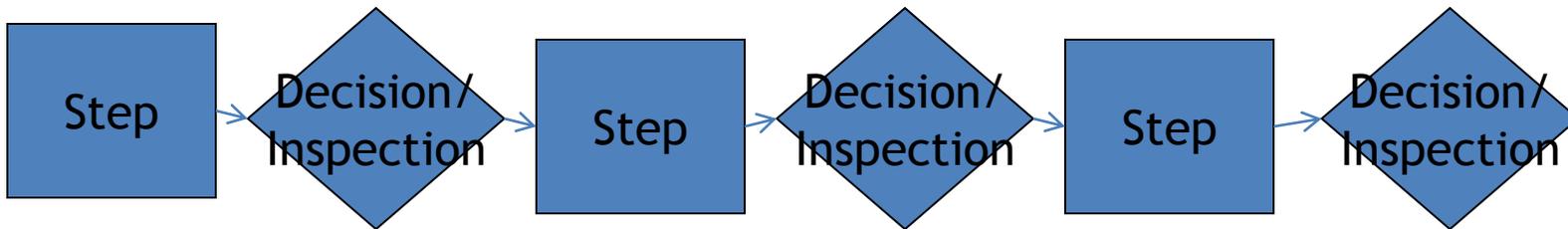
**Binary Logic:**  
Good for analyzing and displaying decision-making processes

# Analysis Tool

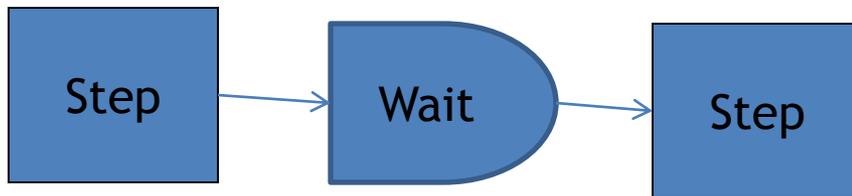
## Things to look for



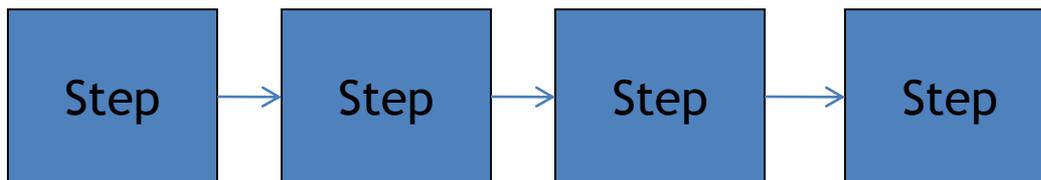
Re-work  
Loops – esp.  
long ones



Multiple  
inspections



Wait states



Too many steps;  
Hand-offs

# Detailed Flow Charts

## What are we looking for?

- Re-work loops
- Redundancy
- Complexity
- Hand-offs
- Delays/wait states
- Dead ends/orphans
- Decision making/inspection

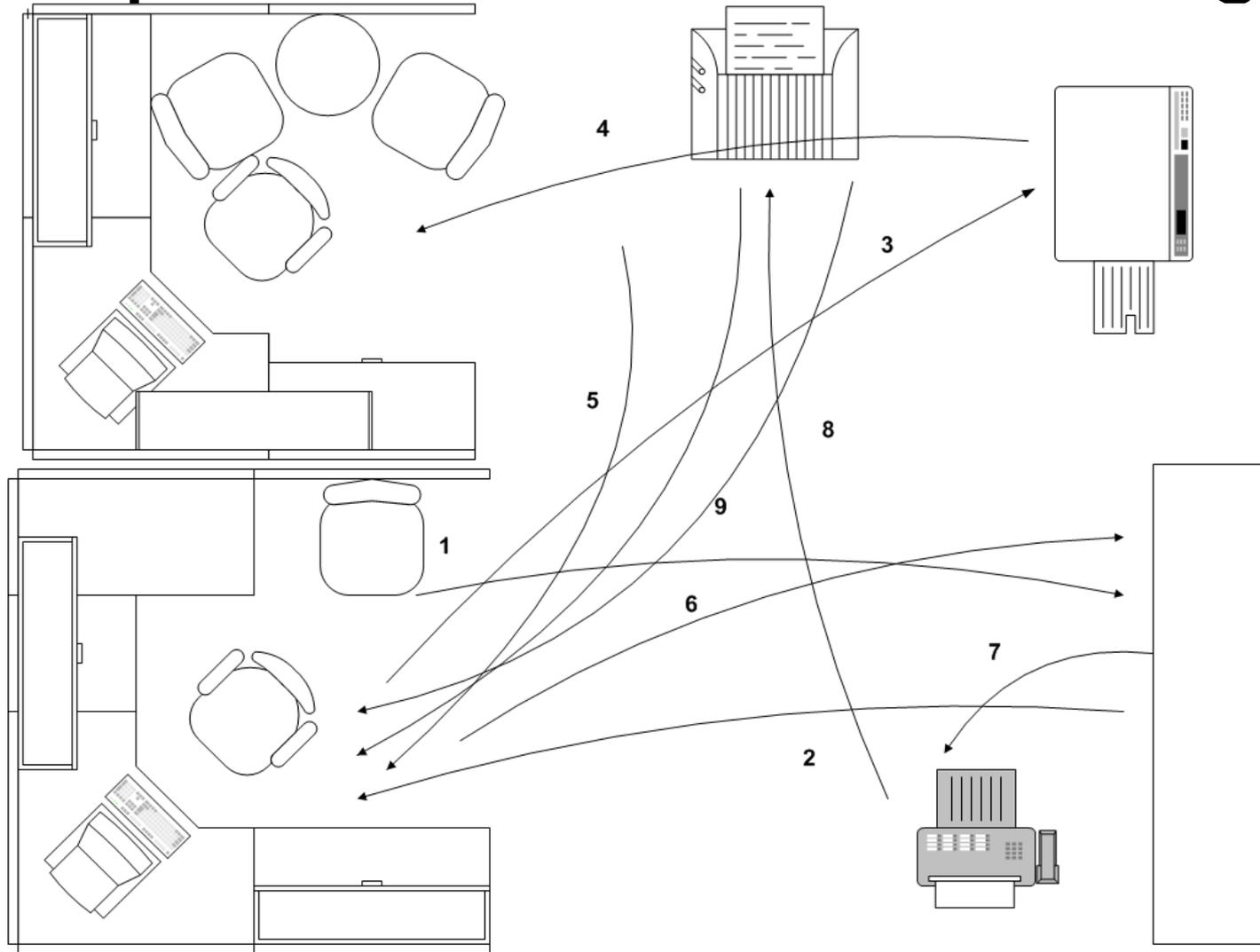
# Analysis Tool

## Detailed Process Flow Chart

- Describes every step and decision
- Calls out documents, data bases, etc.
- *Shows reality* of process: rework loops, undefined specifics, redundancy
- Good for detailed analysis (and design)
- Very time consuming to create
- Sometimes leads to documenting more than is necessary
- Can be hard for others to follow/understand

# Spaghetti Diagrams

A particular form of flow charting



# Flow Charting

## Steps

- Define start and end of process
- Complete high-level flow (conceptual) with 4-7 objectives in between
- Add detail to each objective as needed in matrix format
- Detail those portions of the process where data/information point to possible root cause

# Flow Charting Exercise

- Flow chart the “leaving for vacation” process
- Follow the recommended steps
- Use Post-Its
- You have 10 minutes

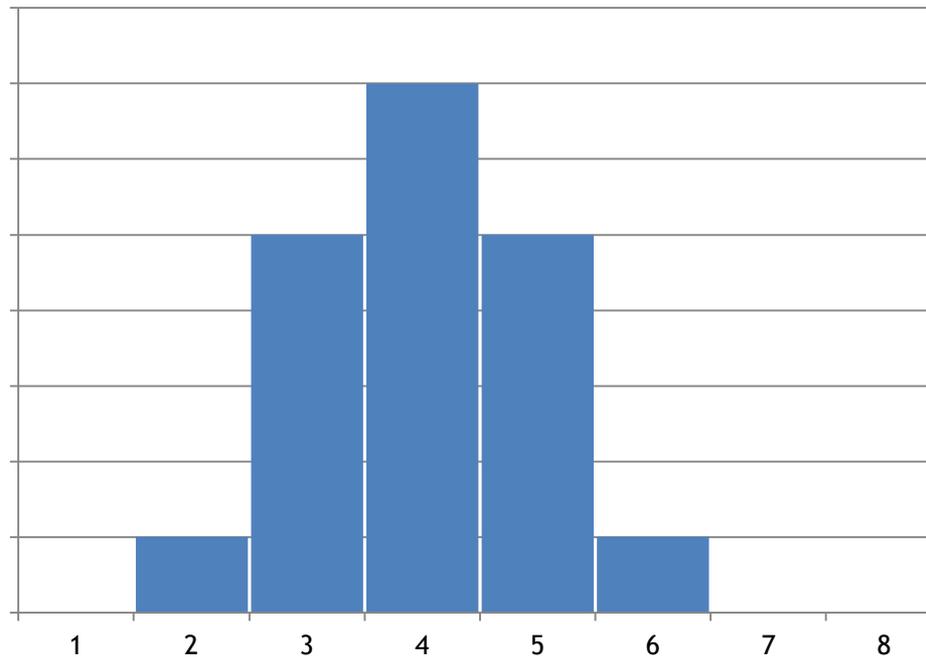
# Flow Charting

## ***Be Careful: Don't get lost in the map***

- The more complex the problem / process, the less useful a detailed process flow may be (at the outset)
- Be clear about what you are flow charting: The way it “is” or the way it “should be”
- Keep focused on your mission. The more you attempt to document with a flow chart, the more you will be tempted to “fix”

# Histograms

aka ... Frequency Distributions



# Histograms

## When would I use this tool?

- You are working on a cycle-time problem or other continuous variable problem
- You want to understand the pattern of variation within a system, process or process step regarding that variable
- You want to quickly see ...

Range

Center

Pattern

# Histograms

## Continuous Variable Examples

Time

Age

Volume

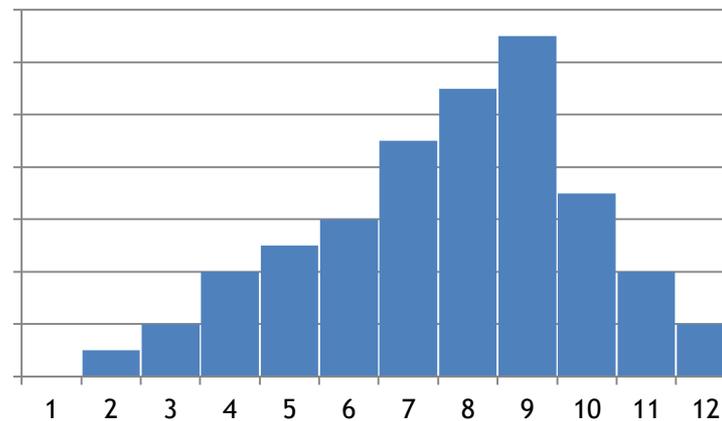
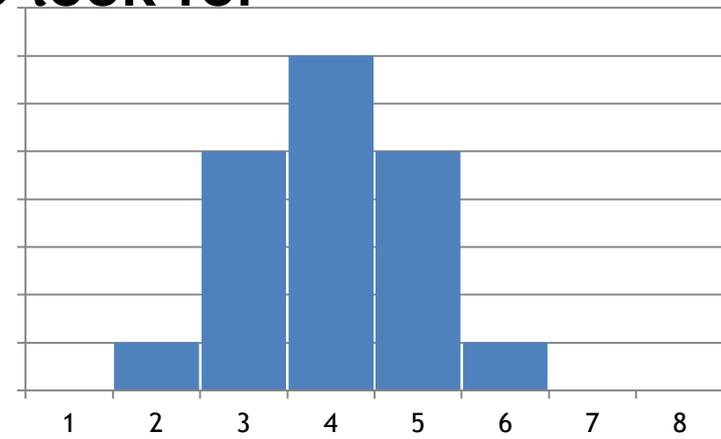
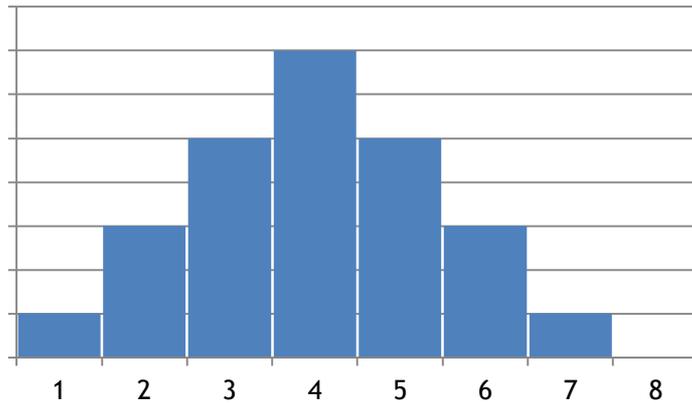
Pressure

Cost

Rates (x per y, etc.)

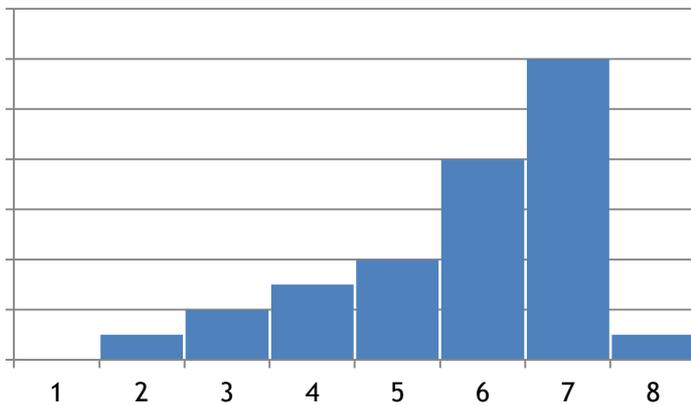
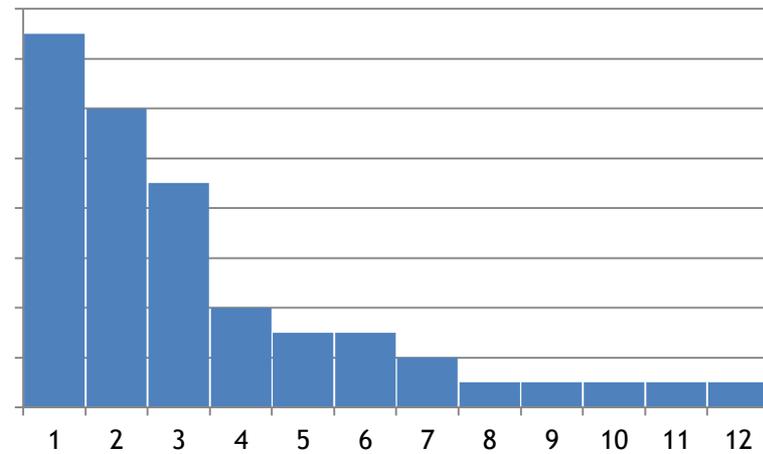
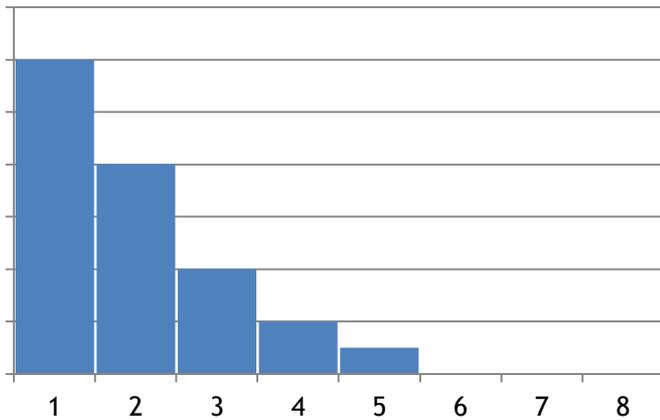
# Histograms

## Patterns to look for



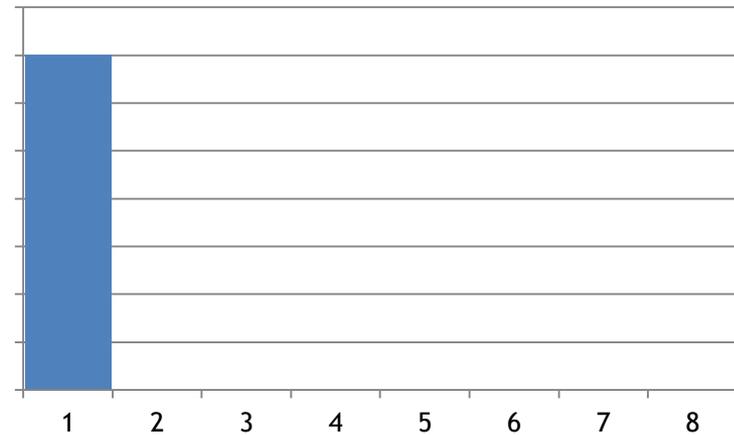
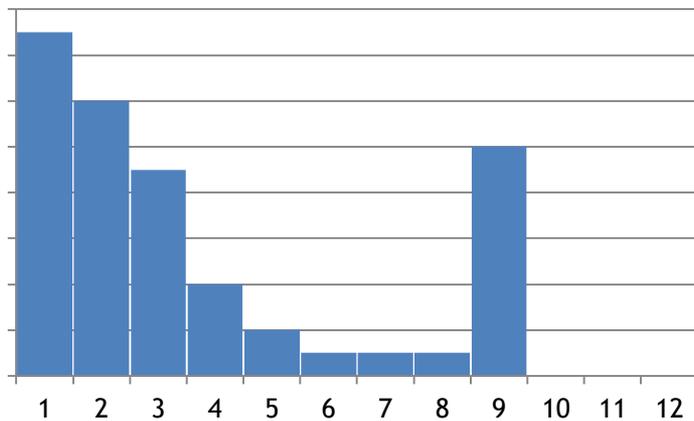
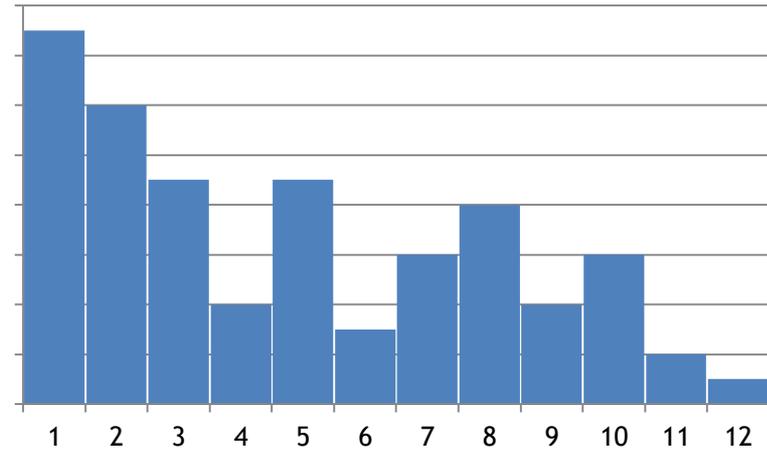
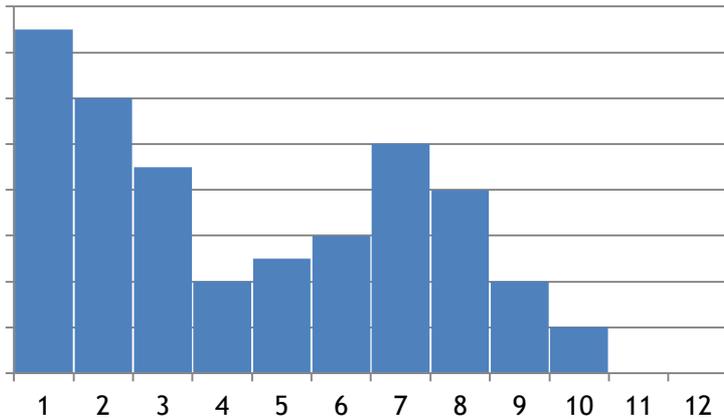
# Histograms

## Patterns to look for



# Histograms

## Patterns to look for



# Histograms

## Interpretation and Next Steps

- Look for ...
  - Biggest contribution to your problem
  - Large ranges
  - Distinct patterns of variation that hint at multiple processes/sources of variation
- Focus on the processes/variables which contribute most to your problem (e.g., most time)
- Focus on the processes/variables which have the most variation (if they contribute to your problem)

# Histograms

## Construction Steps

1. Collect at least 35 data points (and preferably much more)
2. Determine the range of your data
3. Rule of thumb: Divide top range by 7-9
4. Round the dividend to determine roughly how large each interval should be
5. Determine specific intervals, working from low end of range to top of range
6. Make sure intervals don't overlap
7. Plot occurrences within intervals

# Histograms

## Excel Tips

Utilize “histogram” feature ...

1. Enter your data (at least 35 points and preferably much more)
2. Make sure you have Data Analysis Add-On
3. Select Data Analysis/Histogram
4. Create your “bin” ranges (your intervals)
5. Select data range
6. Once frequency distribution is created select those cells and insert a column bar chart

OR...

# Histograms

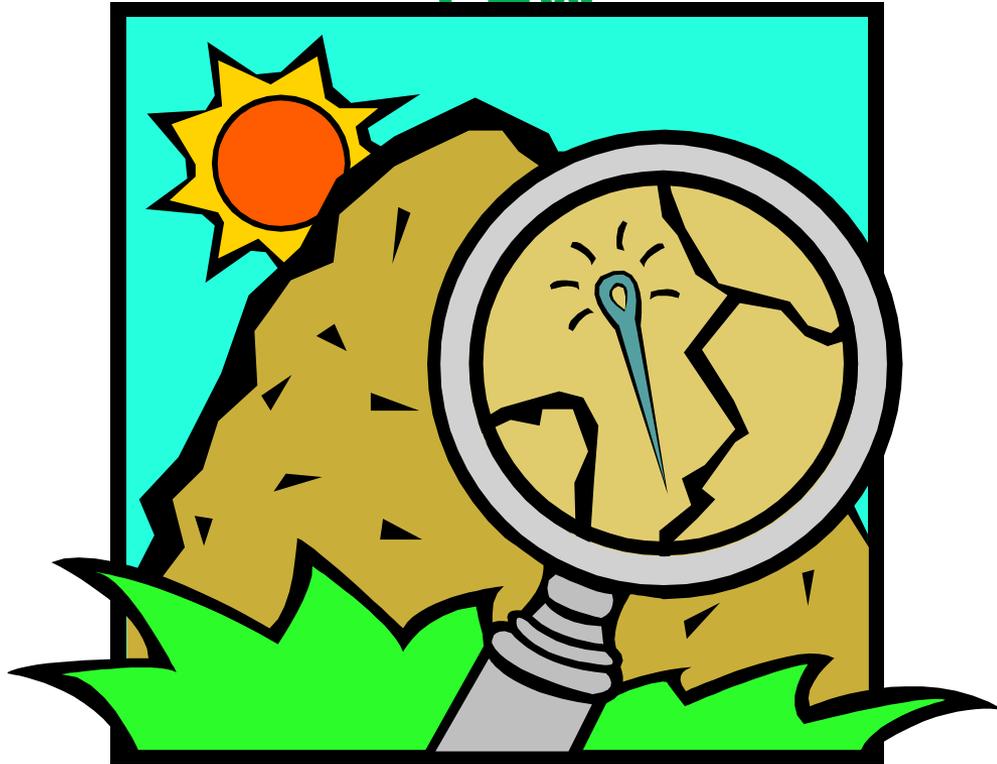
## Excel Tips (2)

Utilize a Pivot Table...

1. Enter data (at least 35 data points)
2. Insert a data table
3. Insert a pivot table
4. Pull data category into “rows,” and pull “count” of that data category into “values”
5. Select a data point under “row labels”, right click and choose “group”
6. Format the group to equal the range of your data and the size of each interval (it’s easy to change so trial and error is ok)
7. Select the chart and insert a column bar chart

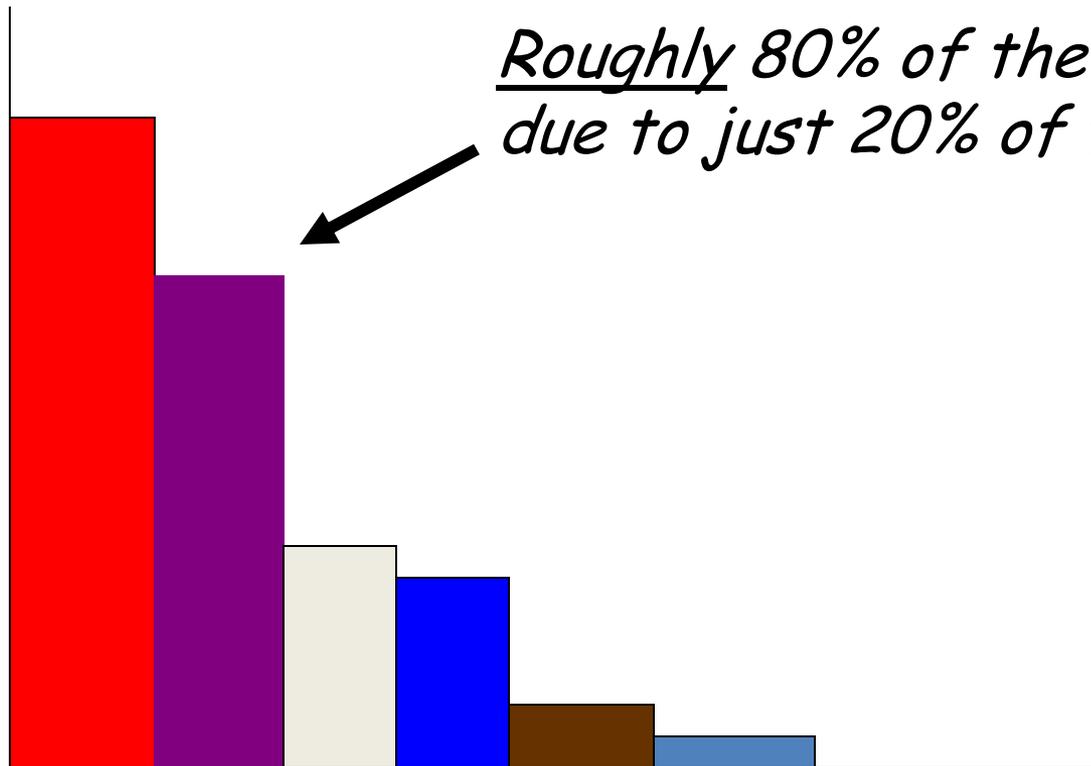
# Pareto Charts & Diagrams

Focusing on the “Vital Few”



# Pareto Principle

## 80-20 Rule



# Pareto Principle

## Examples

- The first two years and last two years of life consume 85% of our health care expenditures
- 15% of cars on the road cause 50% of the auto-related pollution
- 10% of U.S. consumers purchase 43% of all beer

# Pareto Charts & Diagrams

- When would I use this tool?

You are working with attribute data - (defects, incidents, errors, etc.)

Anytime you want to focus attention on largest contributors to a “defect” problem

Assess

Analyze (early or late)

Evaluate

# Pareto Charts & Diagrams

- Simply listing contributors in descending order of magnitude can ...

Help groups focus where their efforts will have most effect

Help groups let go of lesser issues

# Where is the opportunity to focus?

## How big an impact may I have?

Contributors	Count
Ann	11
Clare	6
Colleen	0
Connor	37
Donna	54
Foster	29
Grace	6
Henry	78
Julia	140
Patrick	18
Suzanne	178
Tom	43

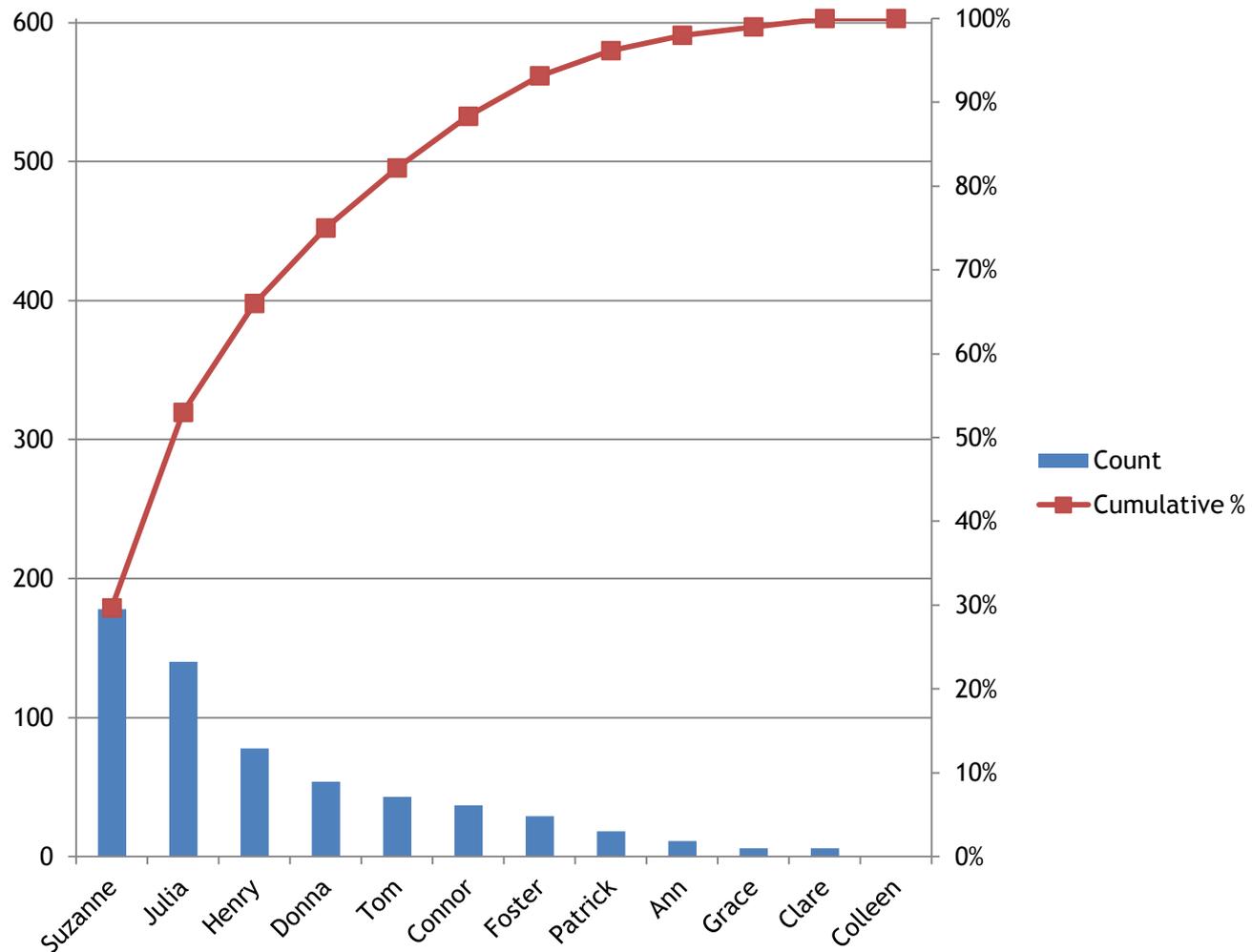
# Pareto Chart

## Example

<b>Contributors</b>	<b>Count</b>	<b>Cumulative %</b>	<b>% of total</b>
Suzanne	178	30%	30%
Julia	140	53%	23%
Henry	78	66%	13%
Donna	54	75%	9%
Tom	43	82%	7%
Connor	37	88%	6%
Foster	29	93%	5%
Patrick	18	96%	3%
Ann	11	98%	2%
Grace	6	99%	1%
Clare	6	100%	1%
Colleen	0	100%	0%
<b>TOTAL</b>	600		

# Pareto Diagram

## Example



# Pareto Chart

## Steps

1. List data (contributors and counts) in descending order of magnitude
2. Calculate total
3. Calculate % contribution to total for each contributor
4. List the cumulative % with each additional contributor

# Pareto Diagram

## Steps

5. Create bar graph with two vertical axes - one for count and the other for cumulative %
6. The left vertical axis should be for the count and the maximum should equal your total count
7. The right vertical axis should be for cumulative % and the maximum should equal 100%
8. The horizontal axis should list each contributor, in order of descending order of magnitude
9. Graph the bars for the count of each contributor
10. Graph the cumulative % for each contributor as a line graph

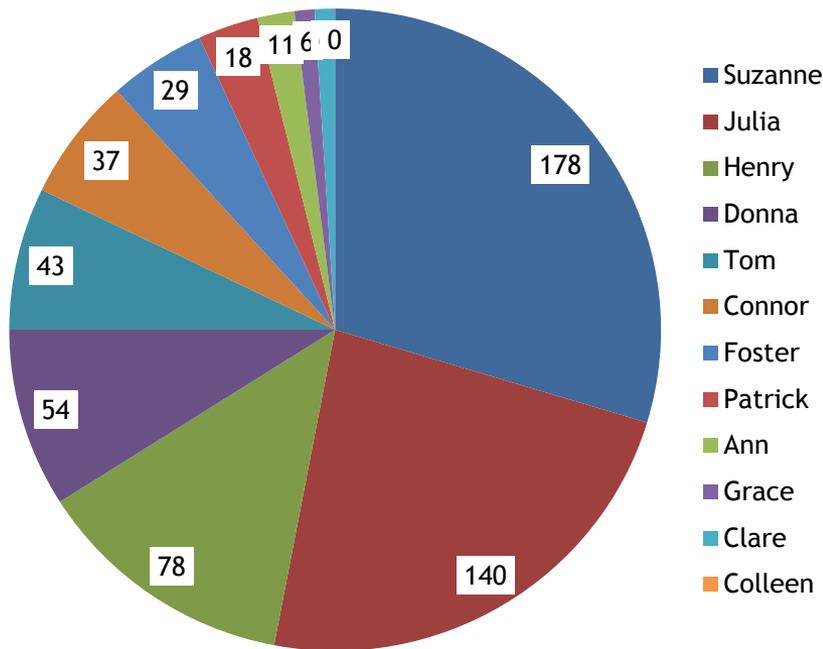
# Pareto Diagram

## Interpreting

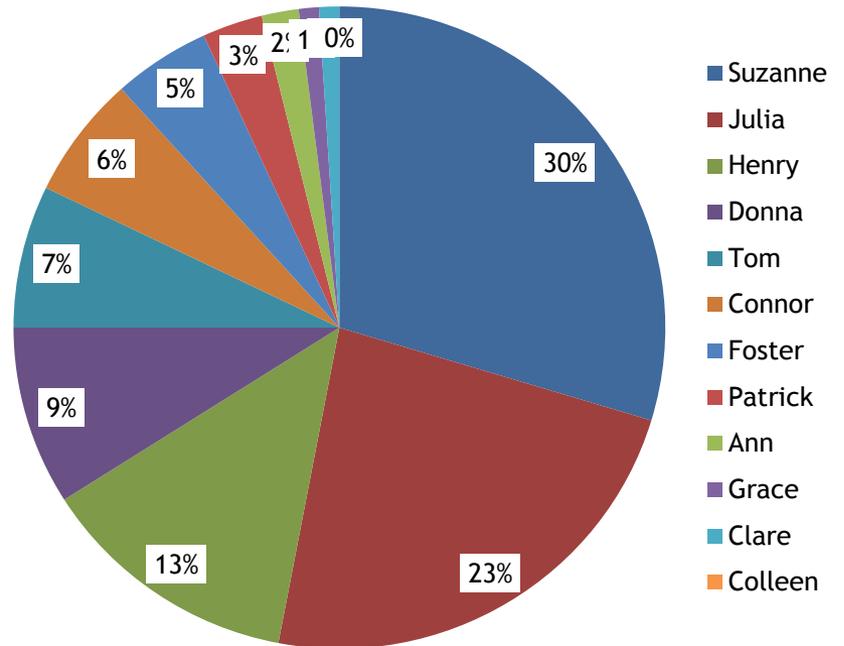
- Start on the left!
- The “vital few” contributors will be the those that comprise 60-85% of total
- You may not have found “cause” at this point
- You may only have identified an area of your problem where further analysis is warranted
- The “useful” many may be addressed later in the project, or in later improvement cycles

# Pie Chart Alternative

Count



% of total



# Box & Whisker Plot



# Box Plot

## When would I use this tool?

- You are analyzing a continuous variable (time, pressure, weight, etc.)
- You want to compare samples side by side on the same scale
- You have *at least* 15 data points per sample

# Box Plot

## What does this tool tell me?

- It gives you a picture of the variation in your data ...

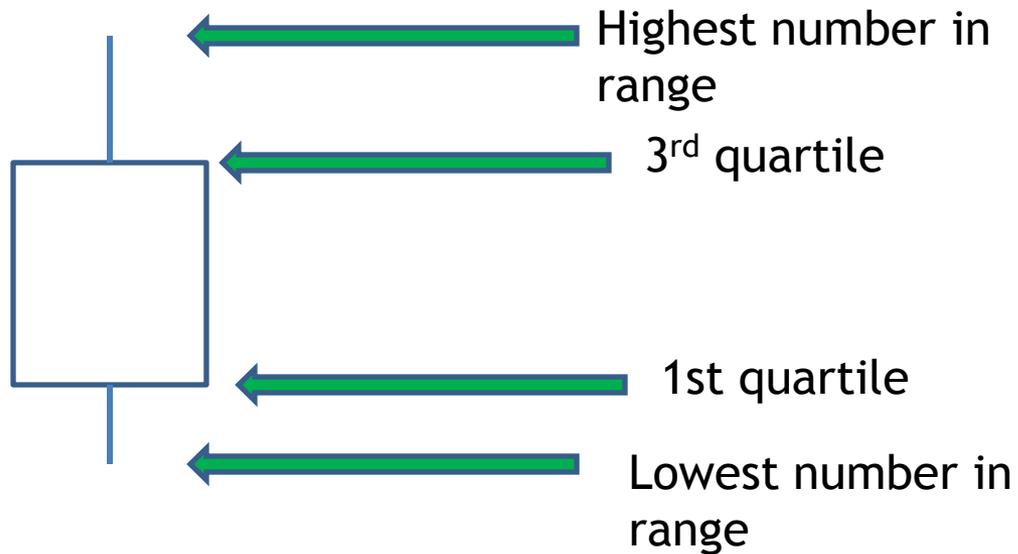
Range (highest, and lowest)

Where most (75%) of your data fall

A sense of the center of your data

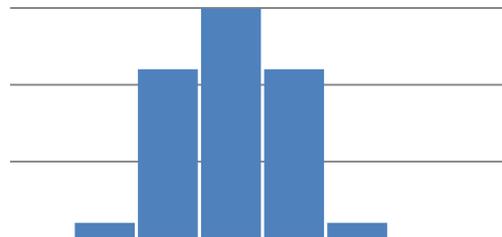
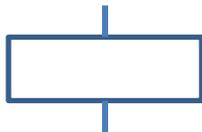
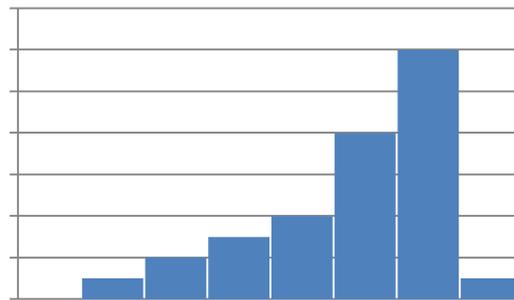
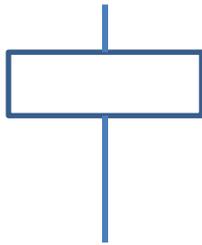
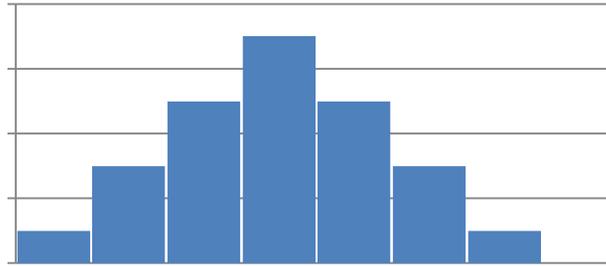
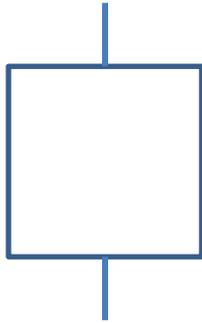
# Box Plot

Continuous  
variable



Sample Category/Segment/Attribute

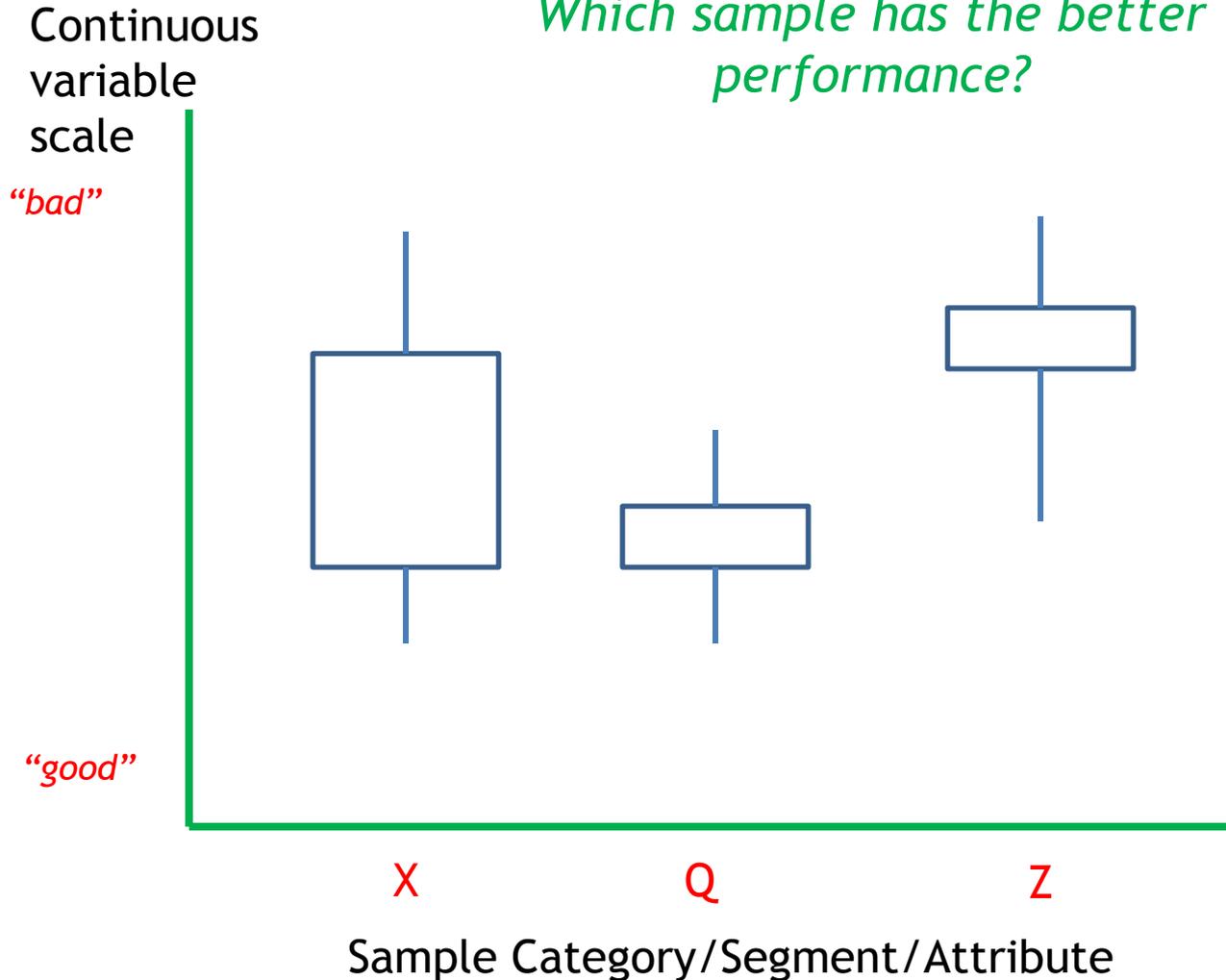
# Box Plots & Histograms



*Box plots provide information similar to a histogram ... but require fewer data points to construct*

# Box Plots

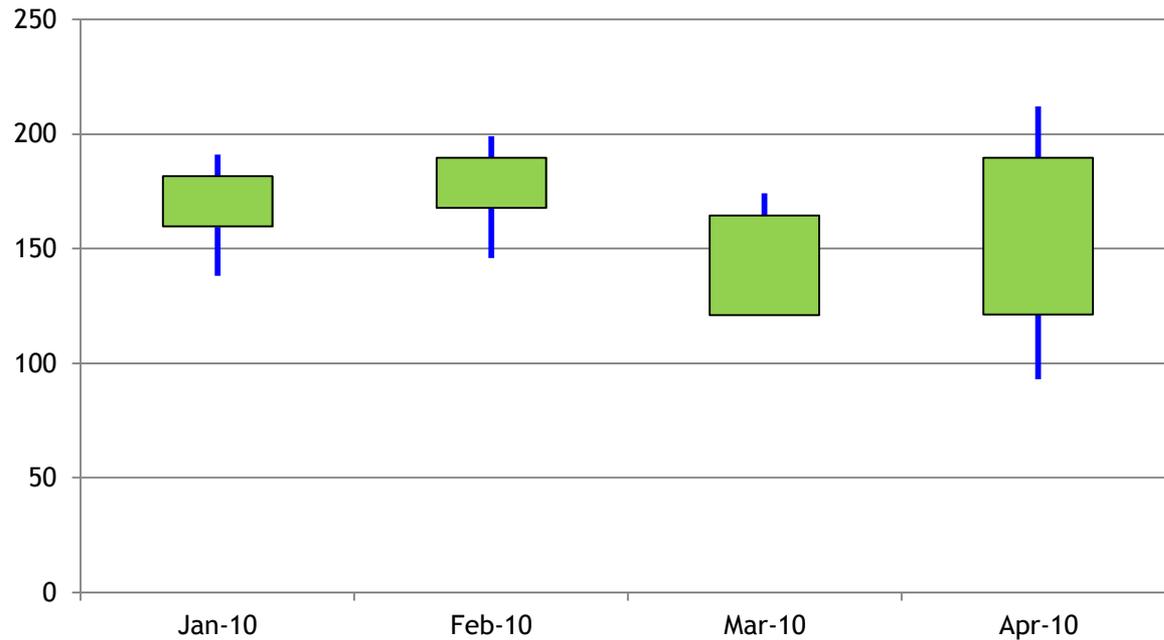
*Which sample has the better performance?*



# Box Plot

## Example - comparing samples over time

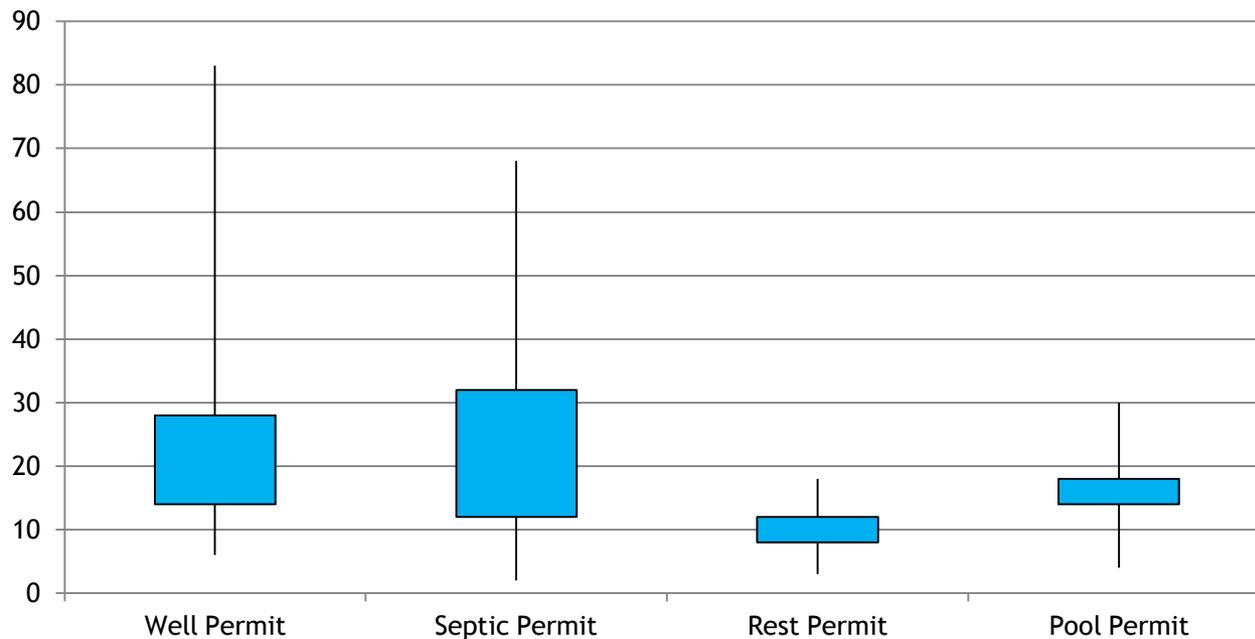
Hours to complete task per unit by month



# Box Plot

Example - comparing samples by different categories

Days to Complete Permit Request  
by permit type



# Box & Whisker Plot

## Steps - by hand

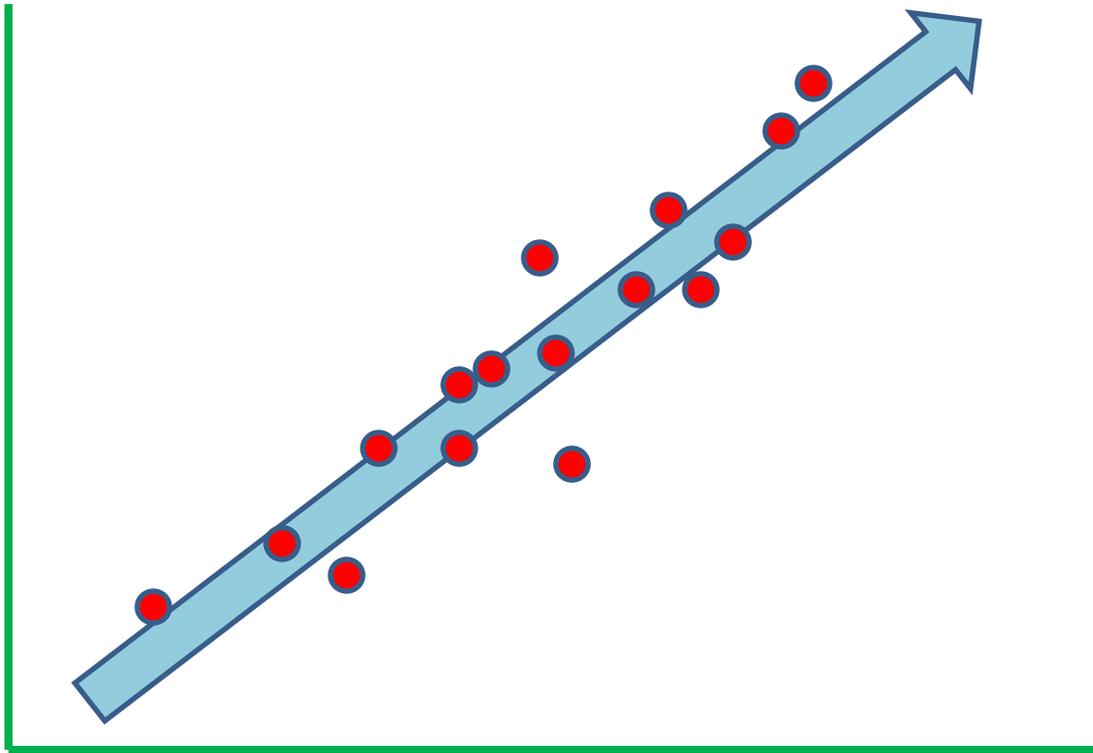
1. Order data from largest to smallest
2. Count the number of data points
3. Determine the highest value
4. Determine the lowest value
5. Determine the halfway point (median)
6. Determine the halfway point of the upper half of data
7. Determine the halfway point of the lower half of data
8. Plot the points

# Box & Whisker Plot

## Steps Using Excel

- Organize data within sample periods/samples
- Calculate Quartiles using Excel Quartile Function[=quartile(xx:xx,0 or 1 or 3 or 4) where 0 = minimum, 1 = quartile, 3 = third and 4 = maximum -- 2 = median]
- Arrange in order of 1st quartile, minimum, maximum, 3rd Quartile
- Apply High-Low-Close Chart from Excel

# Scatter Plot

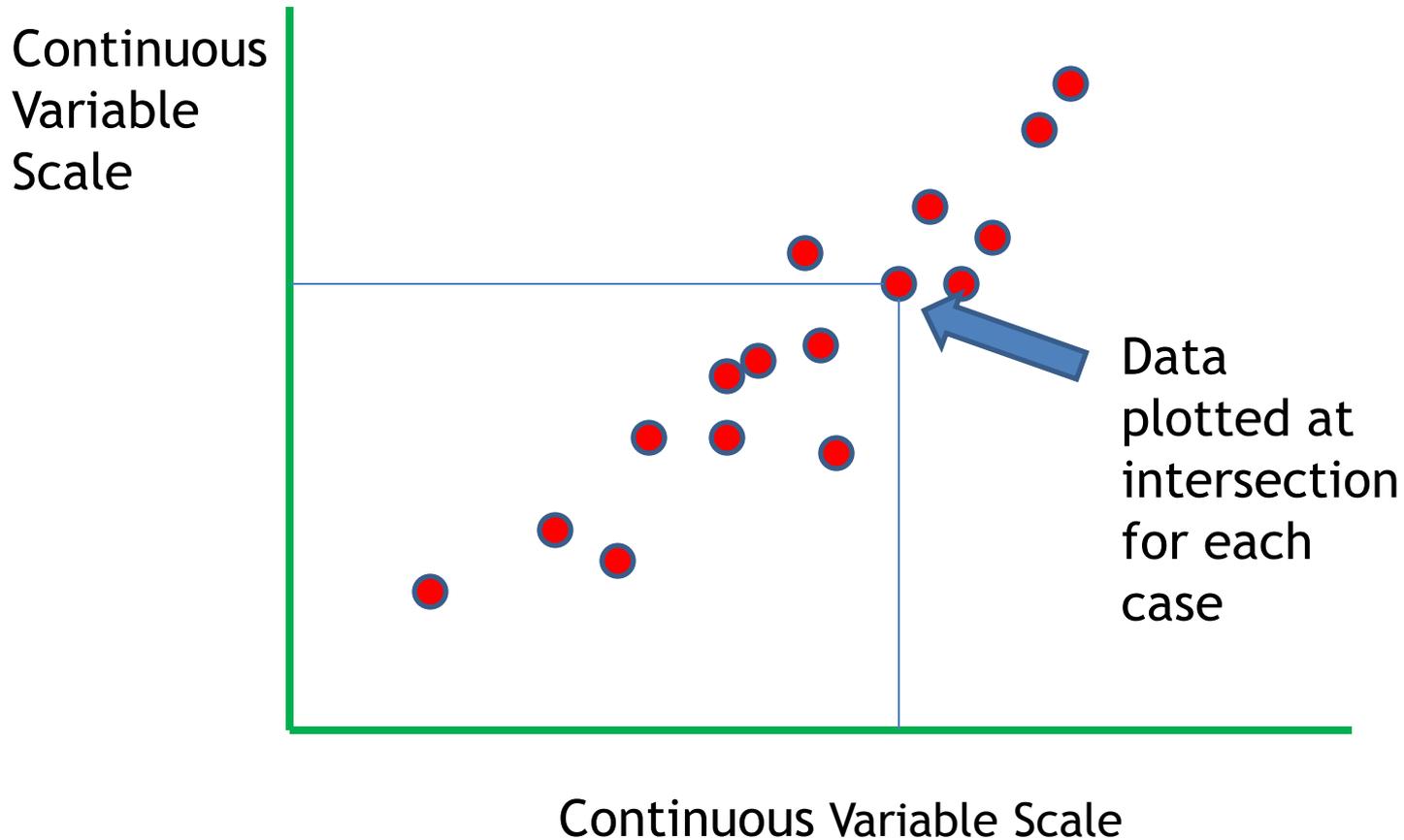


# Scatter Plot

## When would I use this tool?

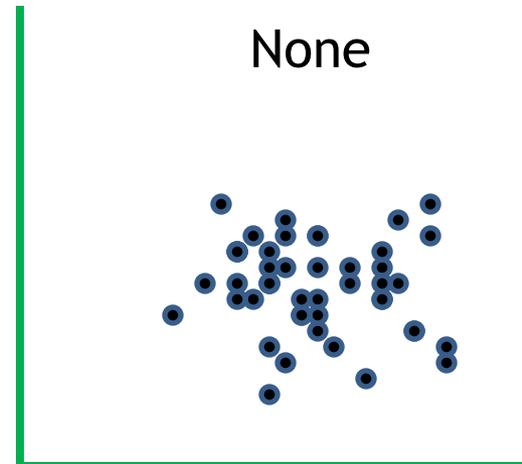
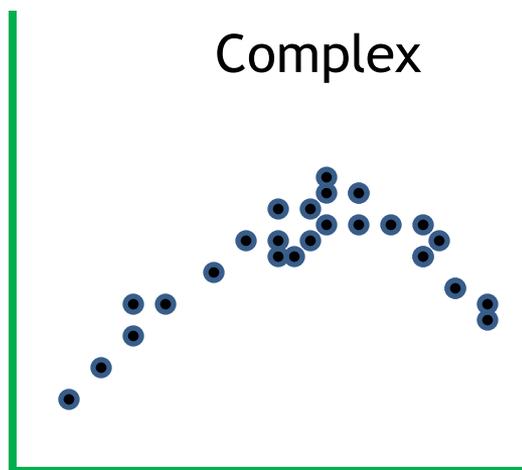
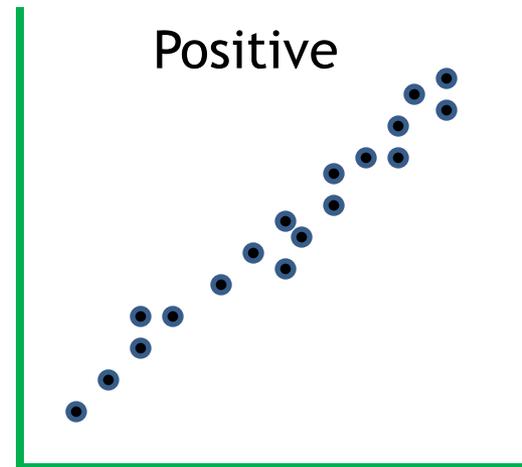
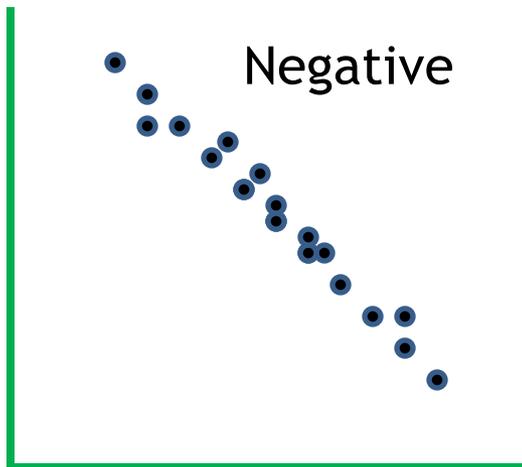
- You want to see if there is a relationship between 2 continuous variables
- You want to test a theory

# Scatter Plot



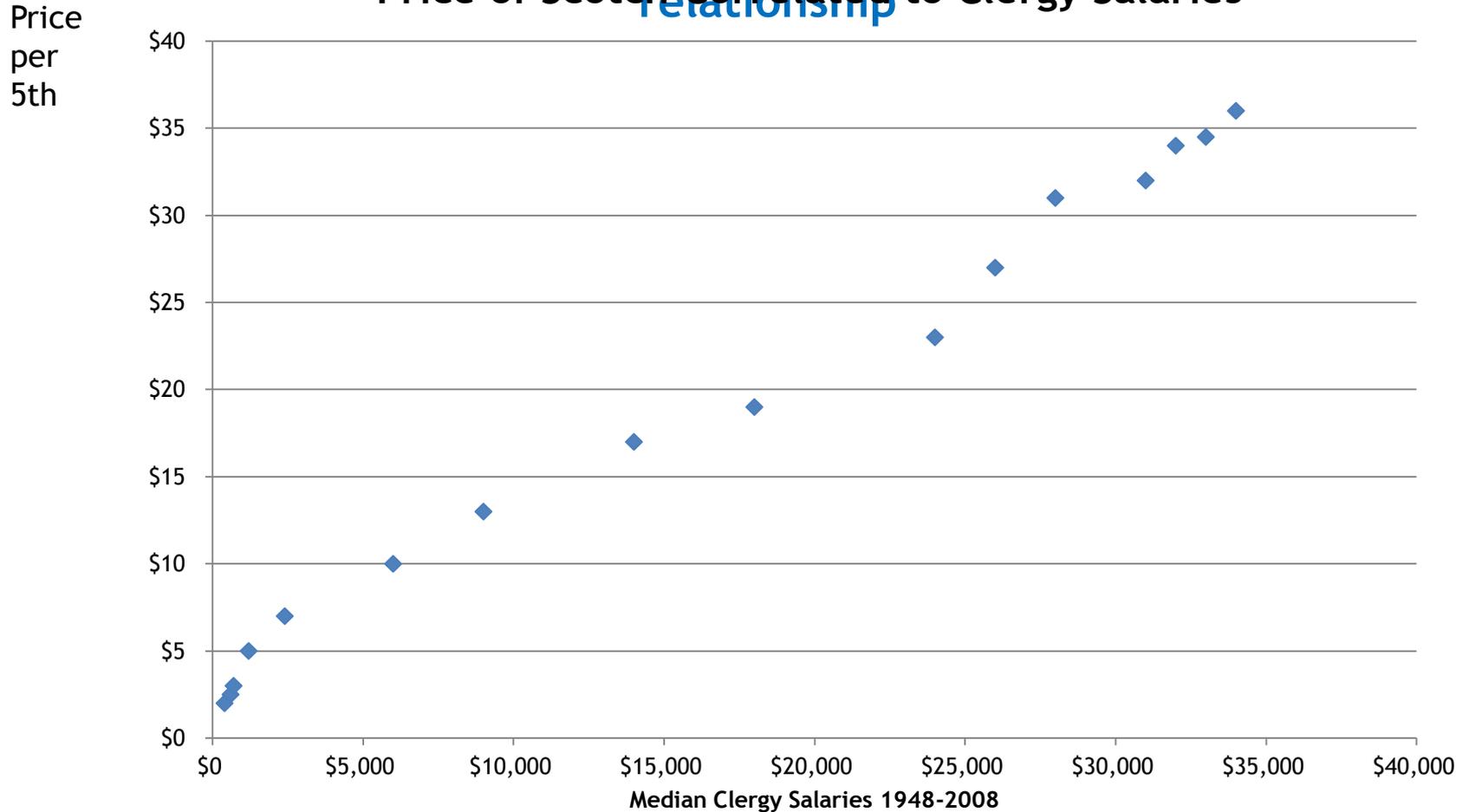
# Scatter Plot

## Types of Relationships



# Scatter Plot

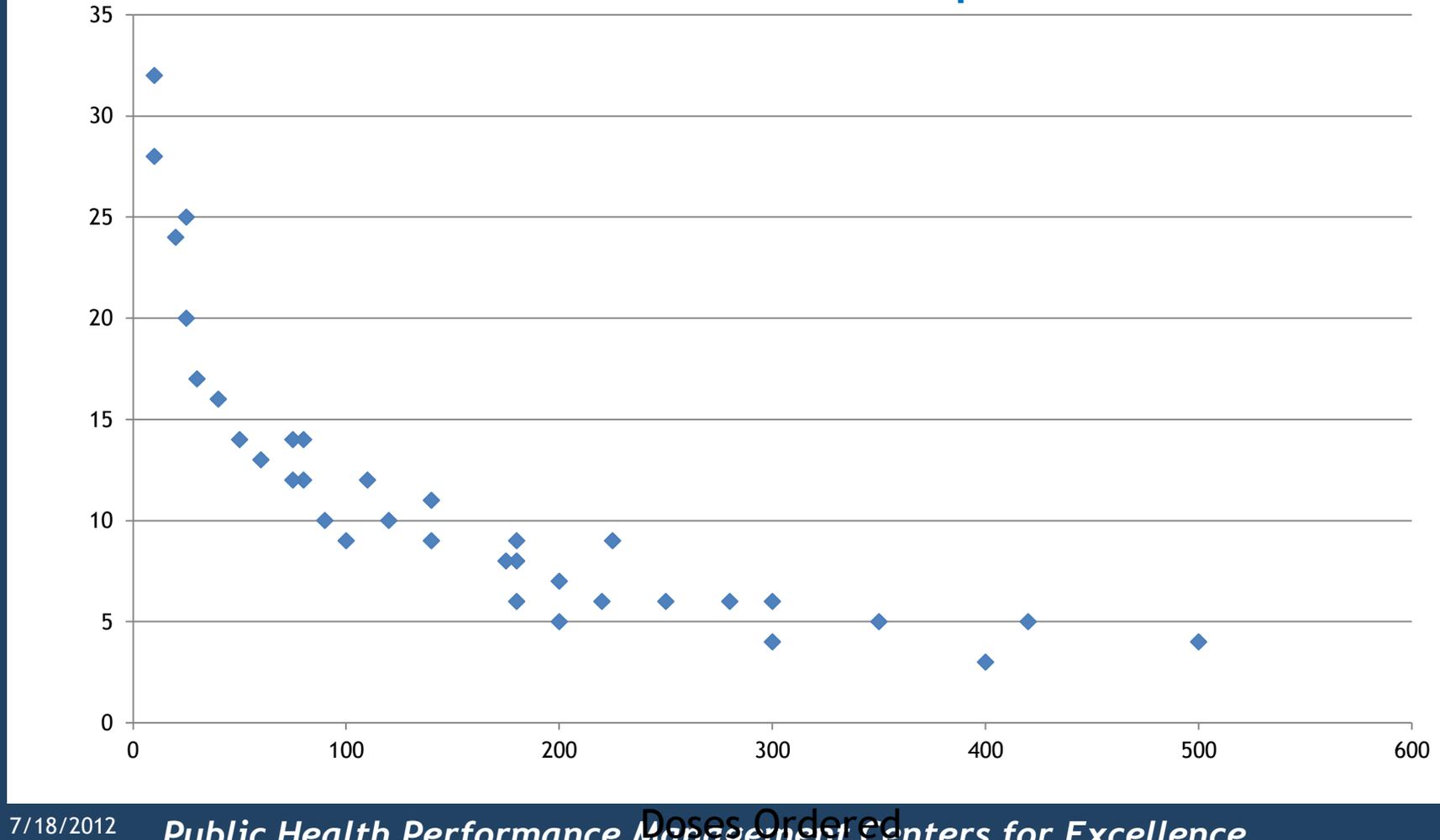
**Example** - Positive correlation without unlikely  
Price of Scotch Correlated to Clergy Salaries



# Scatter Plot

Days to deliver vaccine

**Example** - Negative (complex) correlation with possible causal relationship



# Scatter Plot

## Beware

- Assuming correlation is necessarily a causal relationship
- Over interpreting a vague correlation
- Making inferences with too little data

# Scatter Plot

## Steps - by hand

1. Clarify the relationship you are trying to establish
2. Gather at least 15 data points and preferably much more
3. Create axes for each variable with proportional scales
4. The variable you assume to be causal will go on the horizontal axis (generally)
5. Plot data

# Scatter Plot

## Steps - using Excel

1. Clarify the relationship you are trying to establish
2. Gather at least 15 data points and preferably much more
3. Enter data in side by side columns
4. Choose scatter graph

# Cause & Effect Diagrams

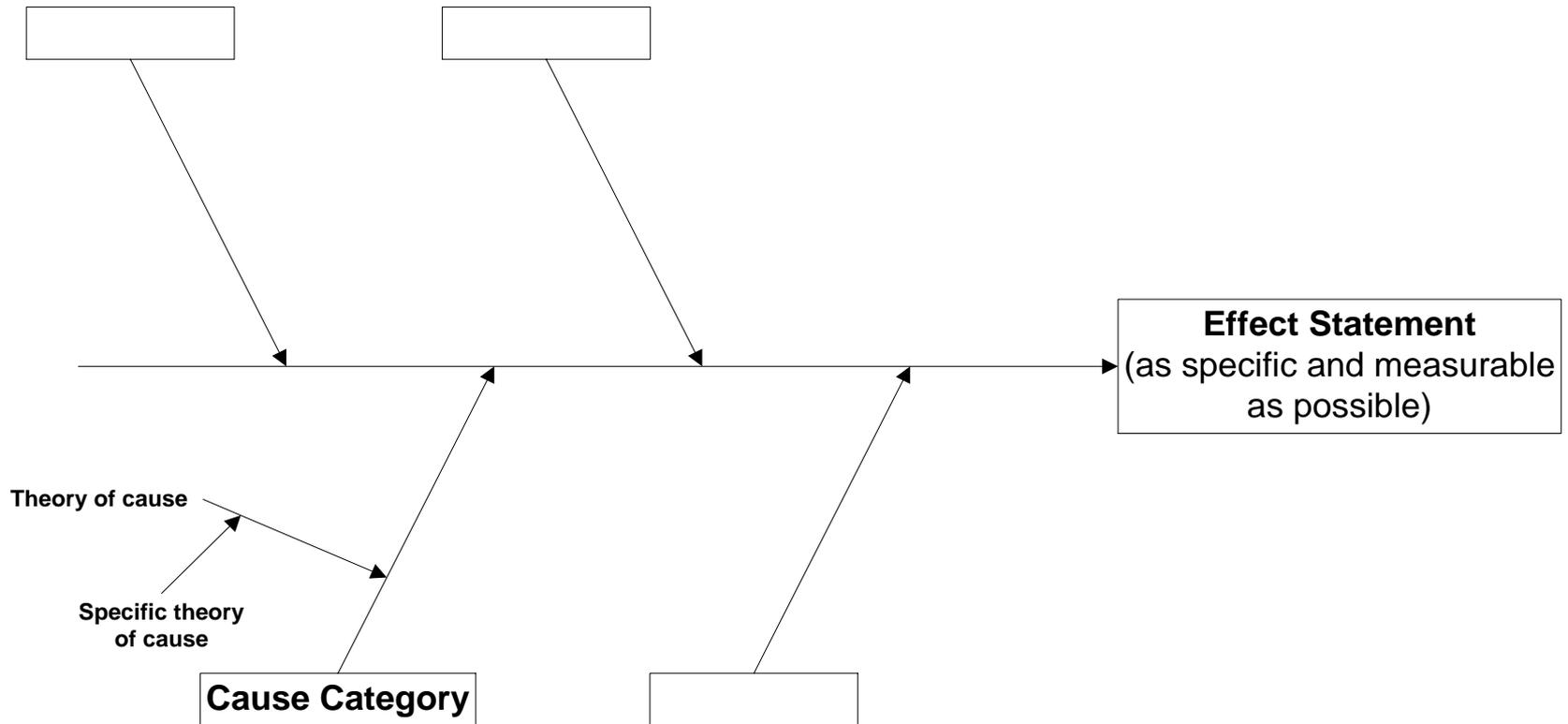


# CEOs

## *When would I use them?*

- Anytime you want to help a team diverge about possible contributors to a problem
- Early in analysis: to help team come up with stratification variables
- Late in analysis: to narrow to very specific causes
- In Change phase: to help team diverge around possible solutions to a narrow cause
- In Program QI: as a focused suggestion system

# Analysis Tool: Cause & Effect Diagram



**“Fishbone” Format**  
(Aishikawa)

# Analysis Tool: Cause & Effect Diagram

Typical Cause Categories

Machines

Materials

Methods (process)

People (men)

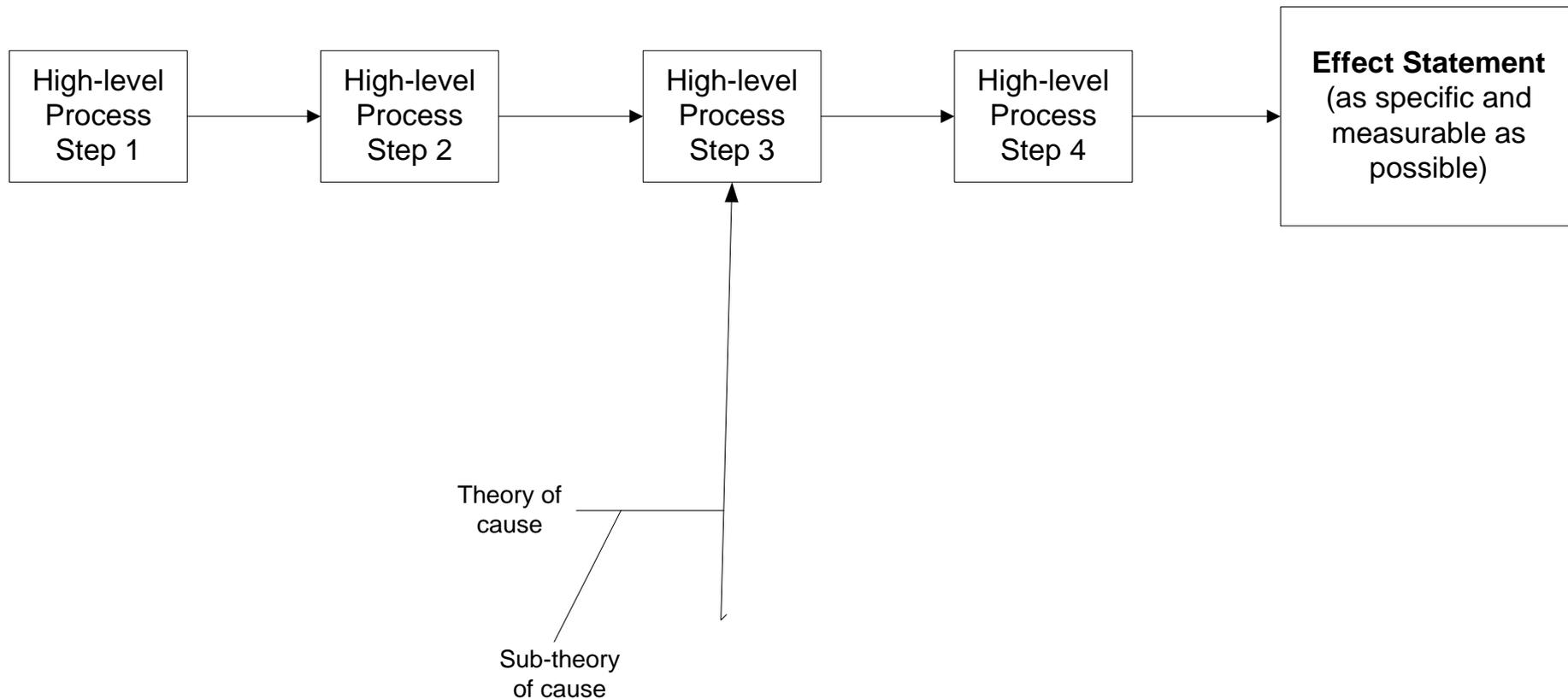
Environment (milieu)

Customer (patrons)

Vendors

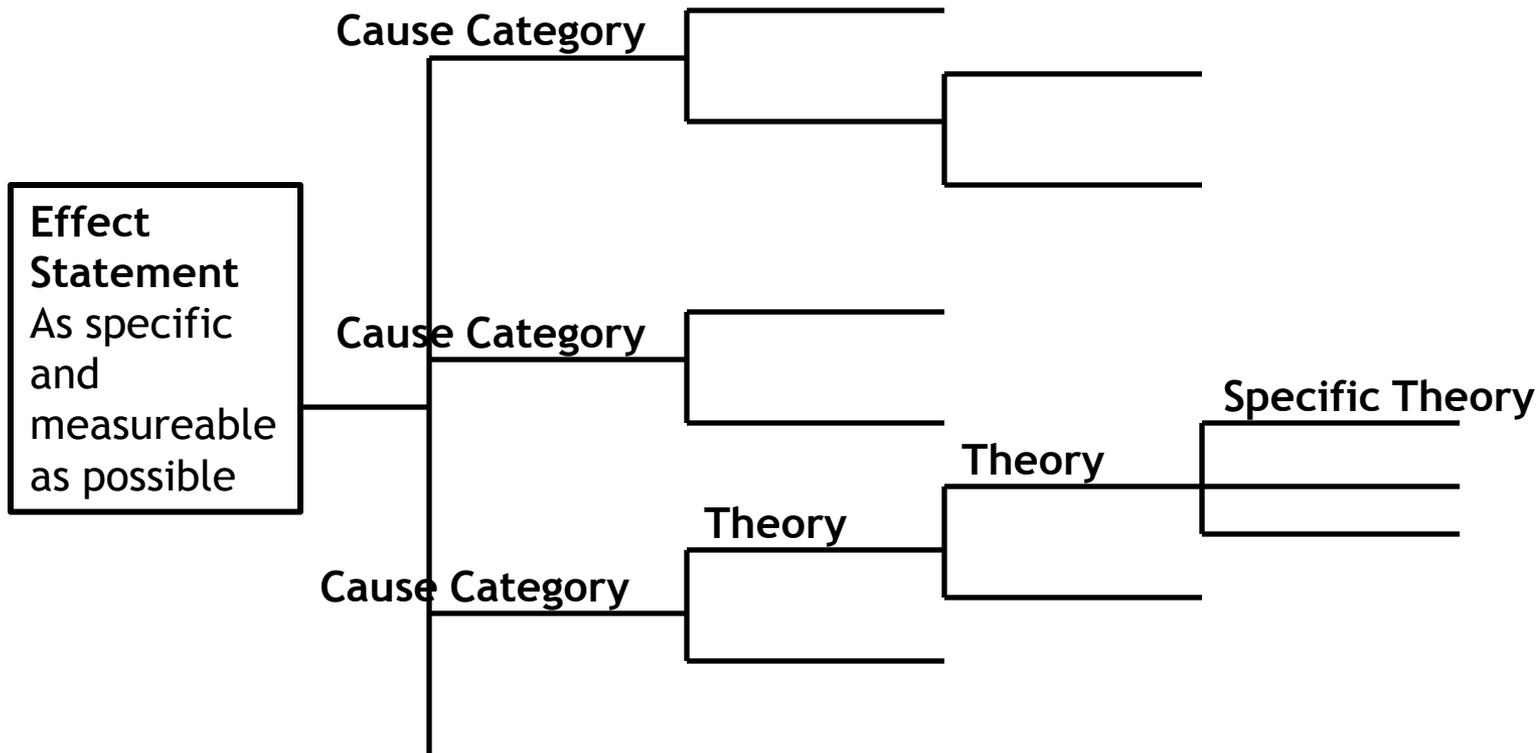
OR: Brainstorm causes on Post-It Notes and then group in natural affinities (see affinity diagram)

# Analysis Tool: CED High-level Flow Format



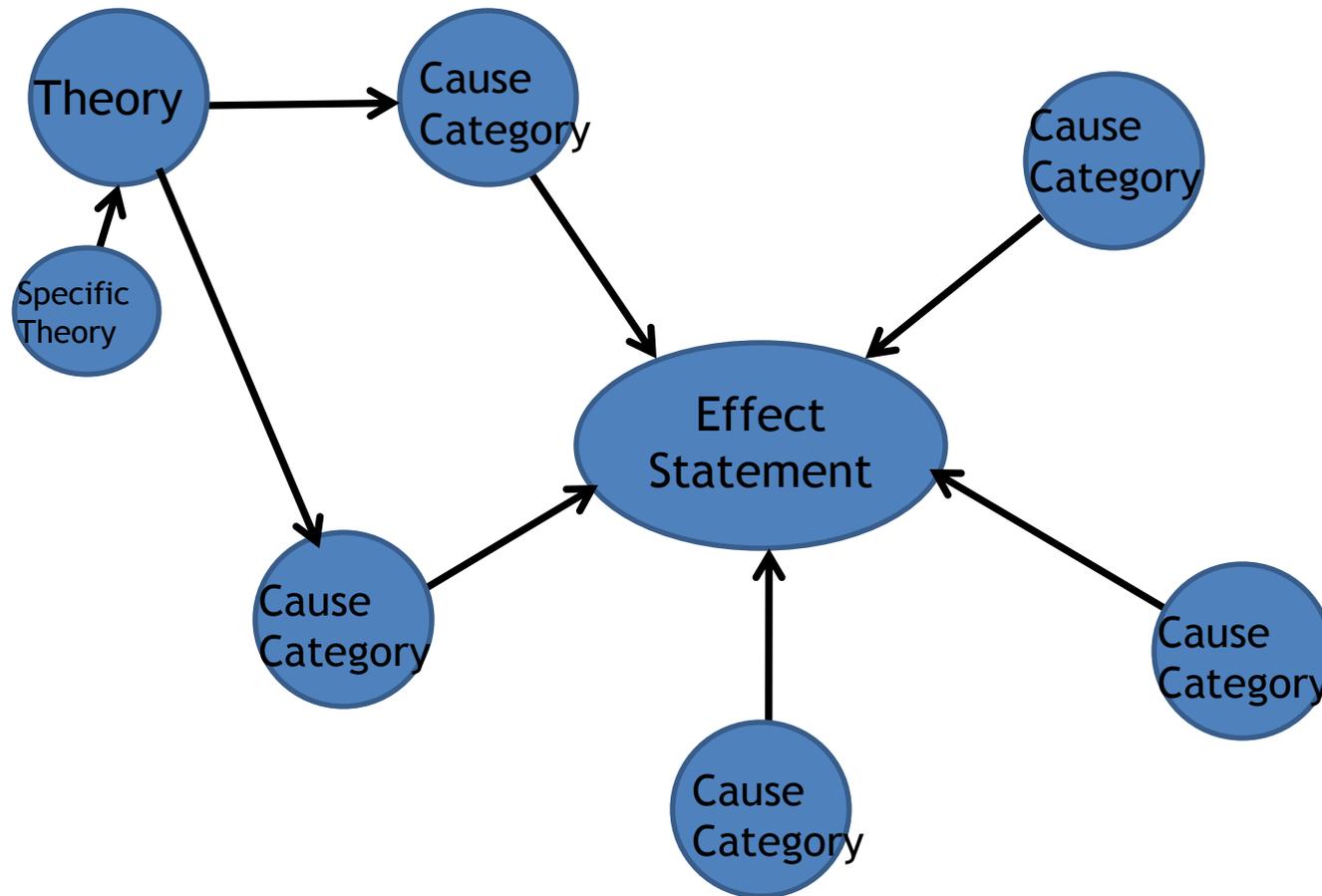
# Analysis Tool: “Why” Root Diagram Format

Why? → Why? → Why? →



# Analysis Tool

## Centric CED Format



# Analysis Tool: Cause & Effect Diagram Steps

- Write specific “effect” statement. Usually, but not always, more specific than your original problem statement.
- Determine format
- Determine cause categories
- Brainstorm theories of cause for each category
- Keep asking “why?”
- When complete, determine top 2-3 causes worth testing/investigating

# Cause & Effect Diagram

## Beware Pitfalls

- Confusing theories with proven causes
- Not “asking why” often enough
- Not diverging enough -- not enough cause categories
- Unbalanced - too much attention to certain cause categories
- Thinking you’re “doing QI” / “evidence based” just because you have a CED