Practical Statistical Problem Solving Using Minitab to Explore the Problem

World Quality Forum
of the International Academy for Quality

Budapest, Hungary

Matthew Barsalou
27 October 2015

SPS

Overview

- Overview
  - Introduction
  - Comparing Differences
  - Investigating Time Related Factors
  - Capability and Stability
  - Exploratory Data Analysis
  - Conclusion

SPS

Introduction

- A problem must be understood before it can be properly addressed.
  - A thorough understanding of the problem is critical when performing a root cause analysis (RCA).
    - An RCA is necessary if an organization wants to implement corrective actions that truly address the root cause of the problem.
  - An RCA may also be necessary for process improvement projects;
    - It is necessary to understand the cause of the current level performance before attempts are made to improve the performance.

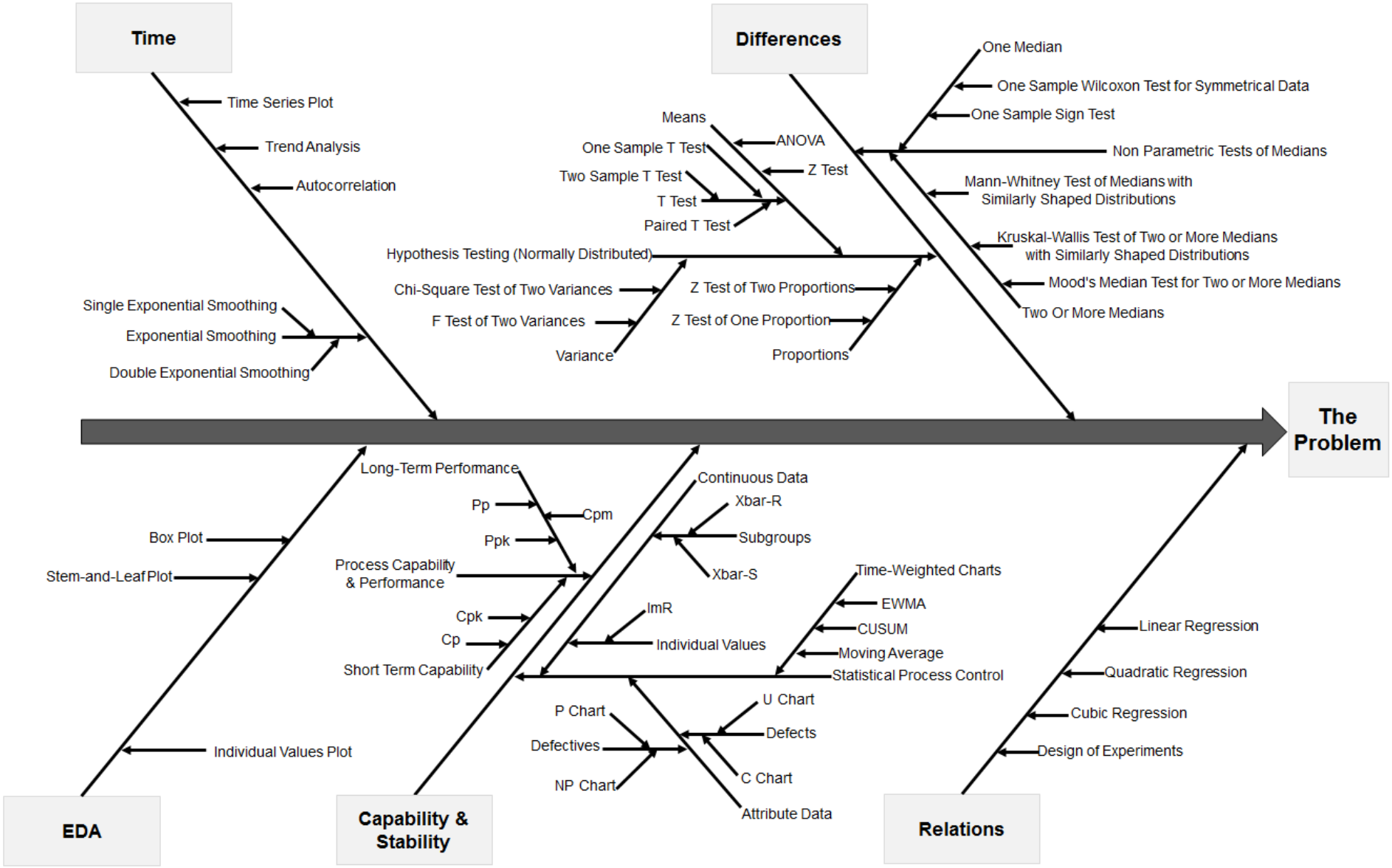
SPS

Introduction

- There are many problem solving related statistical tests that can be performed using the Minitab Statistical Software Program for exploring a problem in the early stages of an investigation.
  - However, the actual test selected should be based upon the type of data and what needs to be understood.
  - Statistical methods can be used to:
    - Explore the problem,
    - gain a better understanding of the problem so improvements can be identified
    - and to monitor performance.

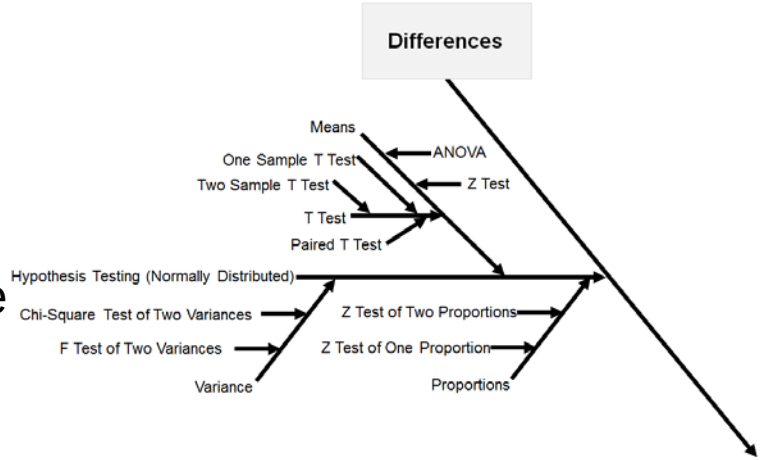
SPS

Introduction



Comparing Differences

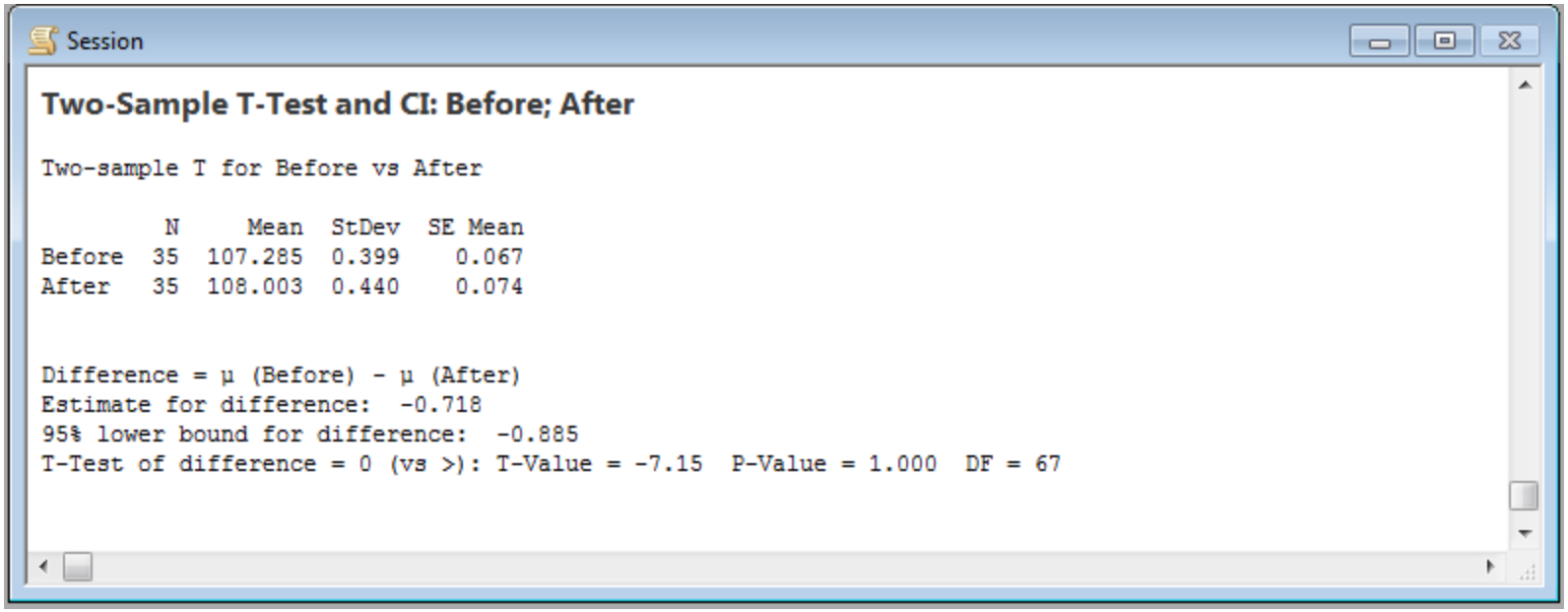
- Hypothesis tests for normally distributed data.
  - Assumes data is normally distributed
  - Can be used to compare to a given value or a second sample:
    - Means
    - Variances
    - Proportions
    - An ANOVA can be performed to compare the means of two or more samples.



SPS

Comparing Differences

- Two sample t test example.
  - A two-sample t-test was performed to ensure that the mean of a critical characteristic improved as a result of improvement actions.

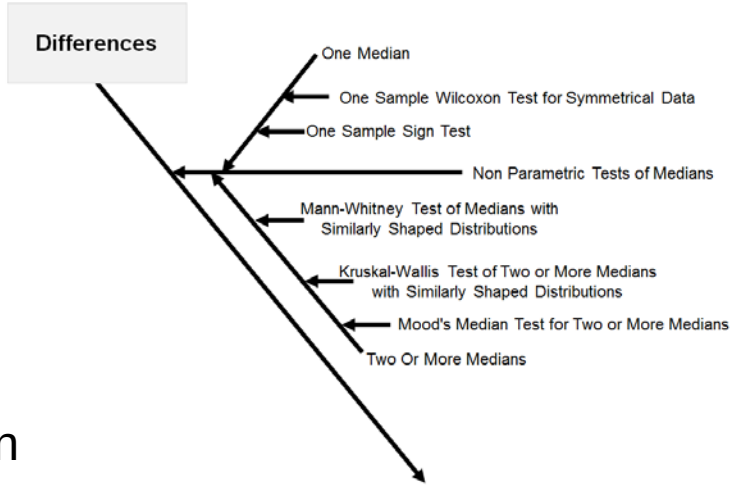


SPS

Examples recreated using representative data

Comparing Differences

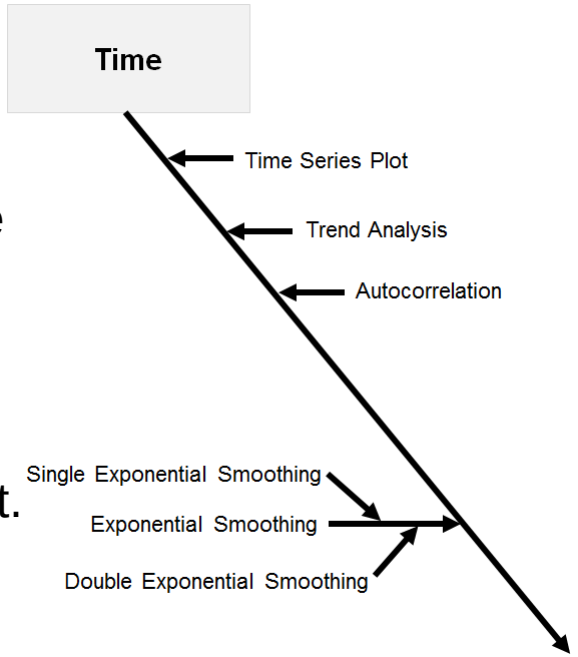
- Non Parametric tests for non-normally distributed data.
  - Used to compare medians.
    - Either to a specified value, or two or more medians.
  - The non-parametric test provide an option when data is too skewed to use other options such as a Z test.



SPS

Investigating Time Related Factors

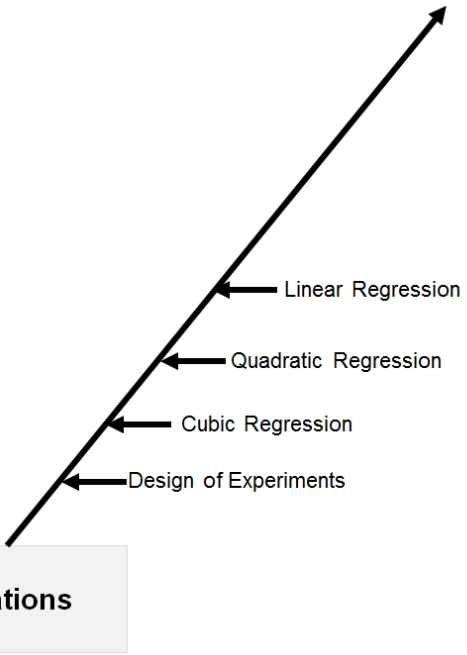
- Time may be of interest when exploring a problem.
  - A time series plot shows each value at the time it was produced.
    - This gives insight into potential changes in a process.
  - A tend analysis is much like the time series plot.
    - Minitab tests for potential tends in the data such as increasing or decreasing values over time.
  - Exponential smoothing options are available to assign exponentially decreasing weights to the values over time when attempting to predict future outcomes.



SPS

Relationships

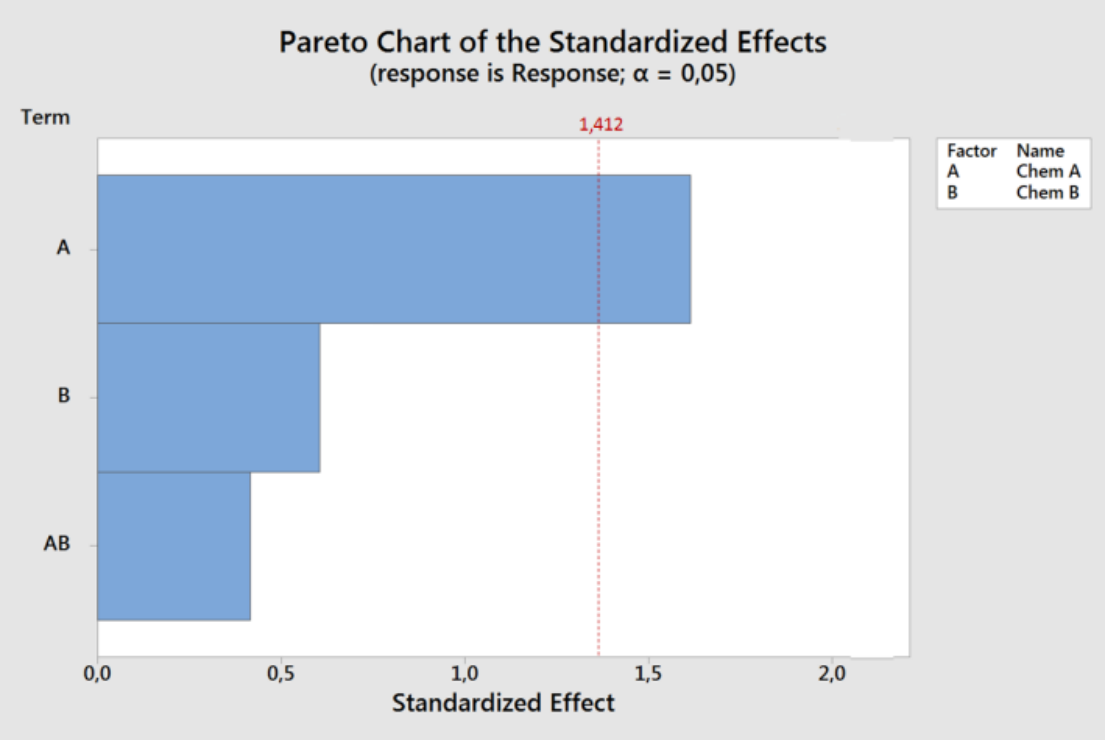
- Relationships can be explored using various types of regression analysis to identify potential correlations in the data.
  - Example: Relationship between the hardness of steel and the quenching time of the steel.
  - This can be helpful when attempting to identify the factors that influence a process.
  - Design of Experiments (DoE) for understanding relationships:
    - Experiments are planed specifically to economically explore the effects and interactions between multiple factors and a response variable.



SPS

Relationships

- DoE example.
  - Cause of aluminum component fatigue failures identified though the use of DoE.

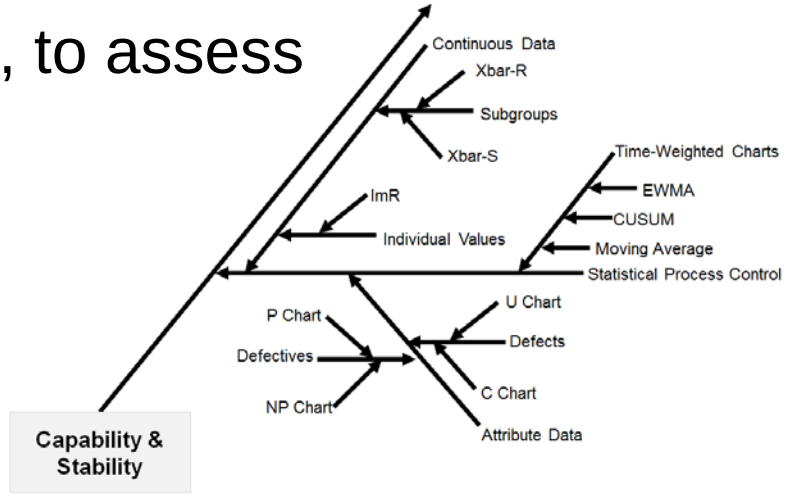


SPS

Examples recreated using representative data

Capability and Stability

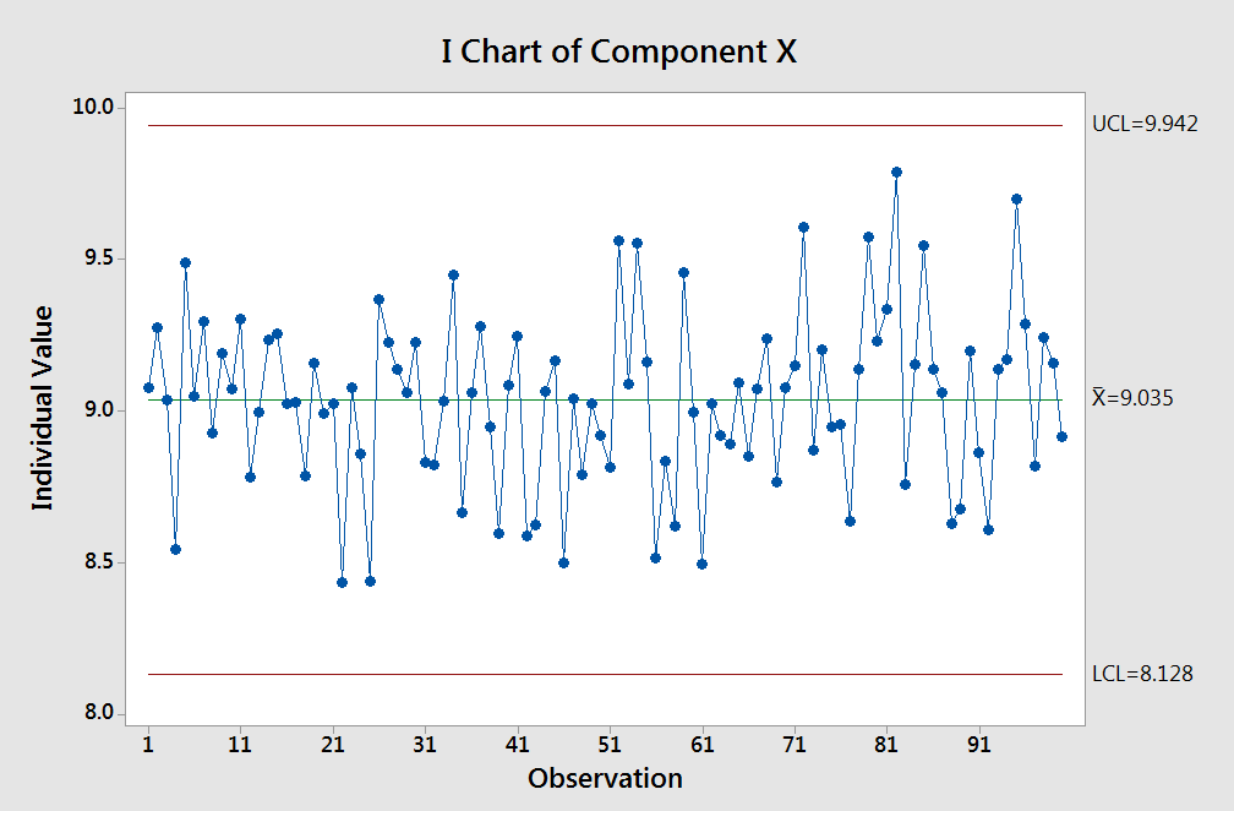
- Statistical Process Control (SPC), to assess the stability of a process.
  - Many types of control charts:
    - For attribute data.
      - Ex.: Defective units/Defectives
    - Continuous data.
      - Ex.: Diameter/Length
    - Time-weighted charts.
      - Values not give equal weights.
  - For assessing the current performance of a process.
    - Determine if the process is in a states of statistical control.
  - For monitoring the performance of a process.
    - After improvements have been implemented.



SPS

Capability and Stability

- Statistical process control example:
  - Use of an ImR chart assured the customer that a component leakage problem was under control and the issue could be closed out.

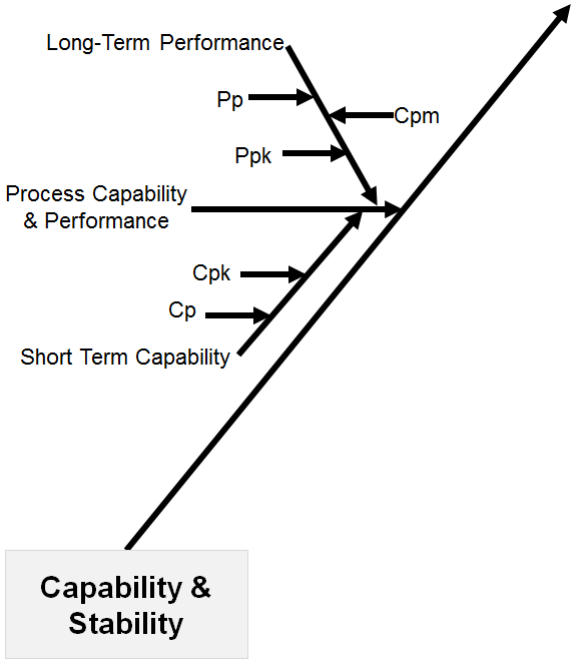


SPS

Examples recreated using representative data

Capability and Stability

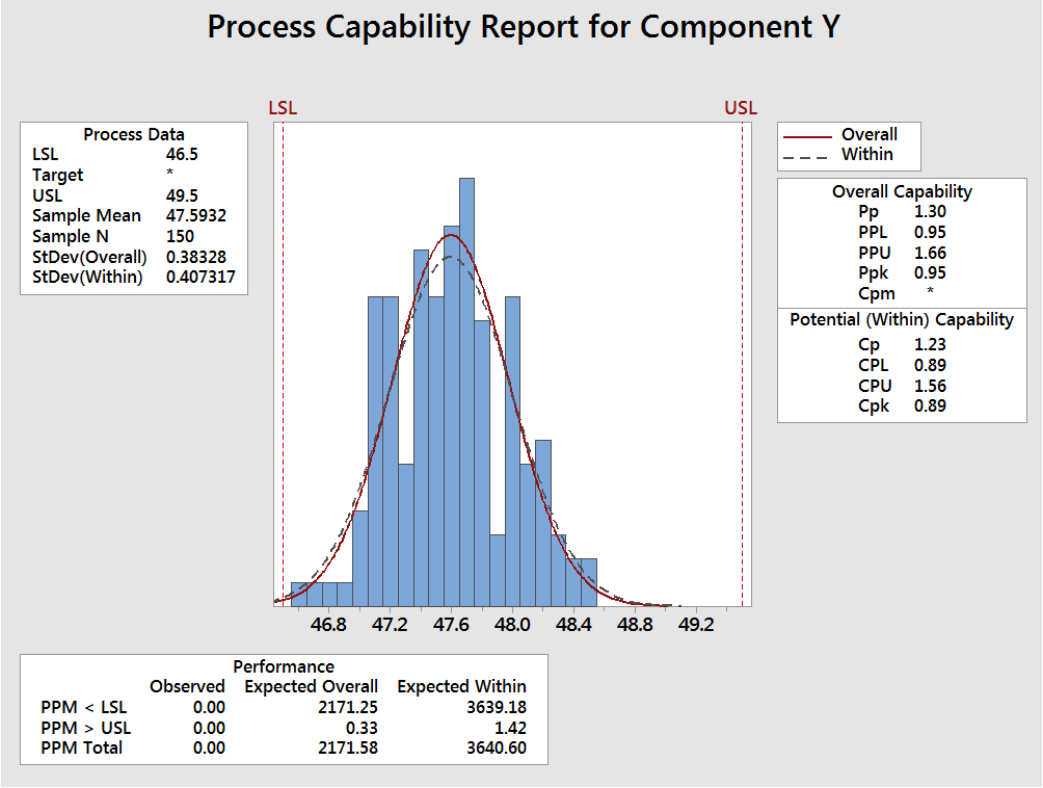
- Measures of process performance and capability can be useful for establishing the baseline performance of a process.
  - This can be helpful in determining of process improvement activities have actually improved the process.



SPS

Capability and Stability

- Capability example:
  - Data from production trial parts were used to determine that there was a potential for out of specification parts.
  - The data indicated process mean needed to be centered.

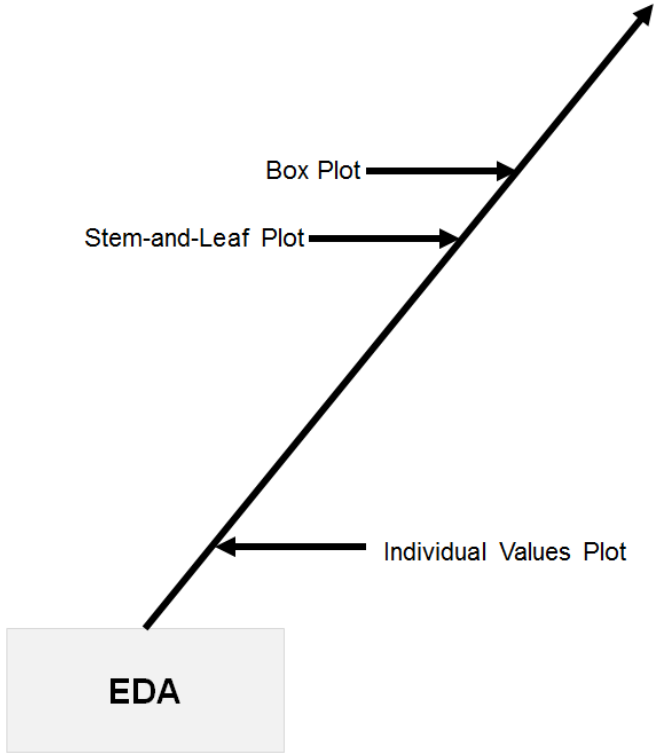


Examples recreated using representative data

SPS

Exploratory Data Analysis

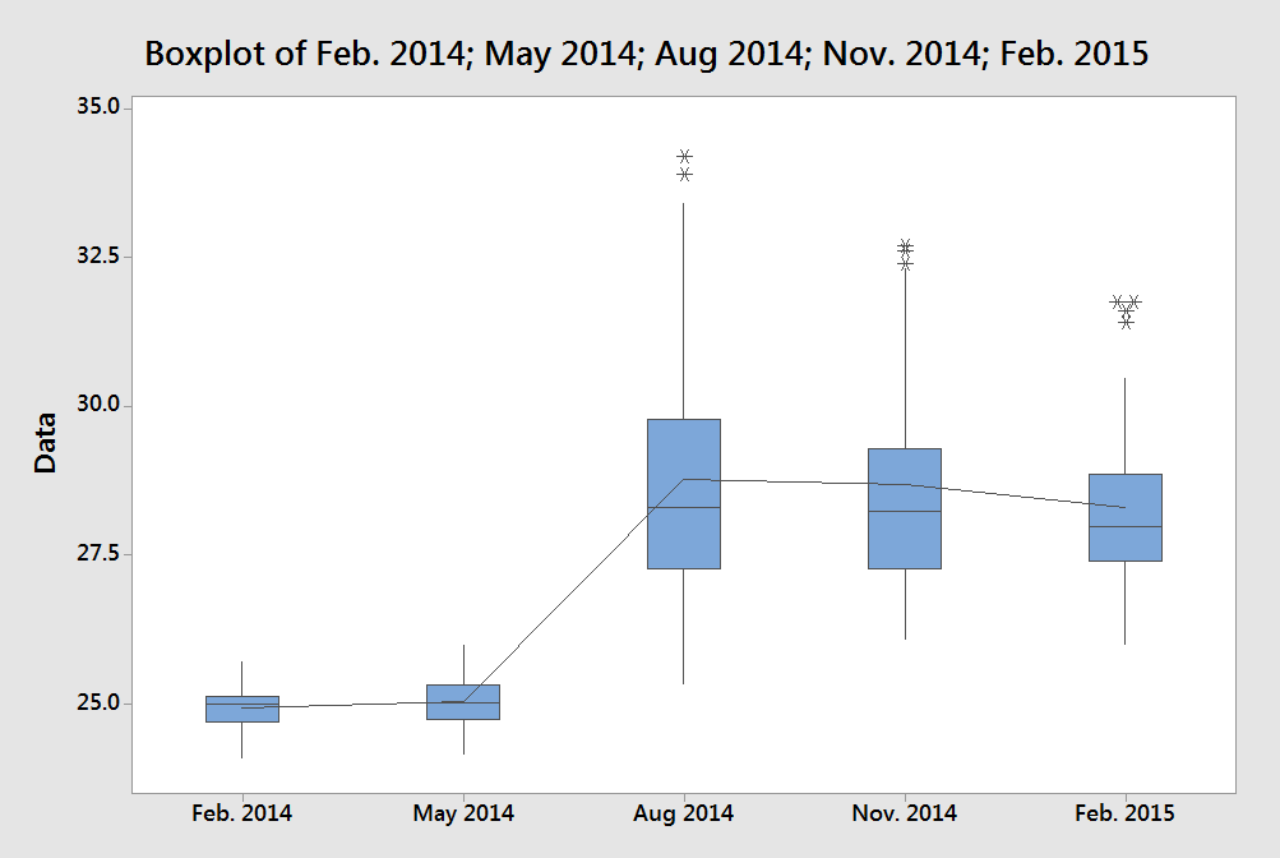
- Exploratory data analysis (EDA) can be useful for gaining insights to the problem using graphical methods.
  - The individual values plot is useful for simply observing the position of each value relative to the other values in a data set.
  - For example, a box plot can be helpful when comparing the means, medians and spread of data from multiple processes.
  - The purpose of EDA is not to form conclusions, but to gain insights that can be helpful in forming tentative hypotheses or in deciding which type of statistical test to perform.



SPS

Exploratory Data Analysis

- Exploratory Data Analysis example:
  - Visualizing the data made it possible to tie changes in process results to casting batches.



Examples recreated using representative data

Conclusion

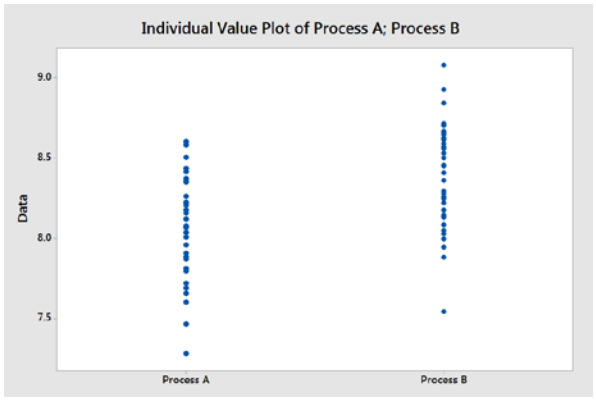
- The tests and methods presented here do not cover all available statistical tests and methods in Minitab.
  - They do provide a large selection of basic options to choose from.
    - These tools and methods are helpful when exploring a problem.
    - Also be helpful for planning and verifying improvements.
  - For example, an individual values plot may indicate one process performs better than a comparable process and this can be confirmed using a two sample t test.
  - The settings of the better process can be used to plan a DoE to identify the optimal settings for the two processes and the improvements can be monitored using an xBar and S chart for the two processes.

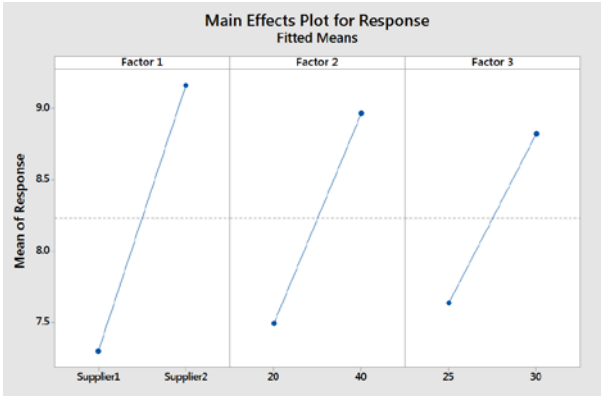
SPS

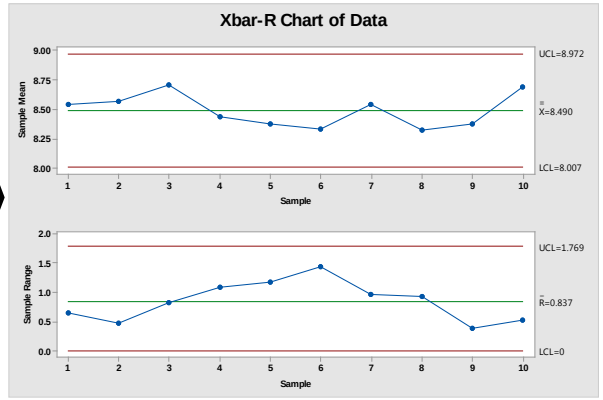
Conclusion

- Hypothetical example:
  - An individual values plot may indicate one process performs better than a comparable process
    - This can be confirmed using a two sample t test.
  - The settings of the better process can be used to plan a DoE to identify the optimal settings
  - The improvements can be monitored using an xBar and R chart for the two processes.

Explore Understand Control







References and Further Reading

- Barsalou, Matthew A. 2015. *Statistics for Six Sigma Black Belts*. Milwaukee, WI: ASQ Quality Press.
- Benbow, Donald W. and T.M. Kubiak. 2009. *The Certified Six Sigma Black Belt Handbook.* Milwaukee, WI: ASQ Quality Press.
- Box, George E.P., Stuart Hunter and William G. Hunter. 2005. *Statistics for Experimenters: An Introduction to Design, Data Analysis and Model Building.* (2nd ed.), Hoboken, NJ: John Wiley & Sons.
- de Mast, Jeroen and Benjamin P. H. Kemper. 2009. “Principles of Exploratory Data Analysis in Problem Solving: What Can We Learn from a Well-Known Case?“ *Quality Engineering.* 21 no. 4: 366-375.
- Lawson, John, Erjavec, John. 2001. *Modern Statistics for Engineering and Quality Improvement.* Pacific grove, CA: Wadsworth Group. 2001.
- Montgomery, Douglas C., George C. Runger and Norma F. Hubele. 2001. *Engineering Statistics* (2nd ed.). New York: John Wiley and Sons.
- Tukey, John W. 1977. *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.