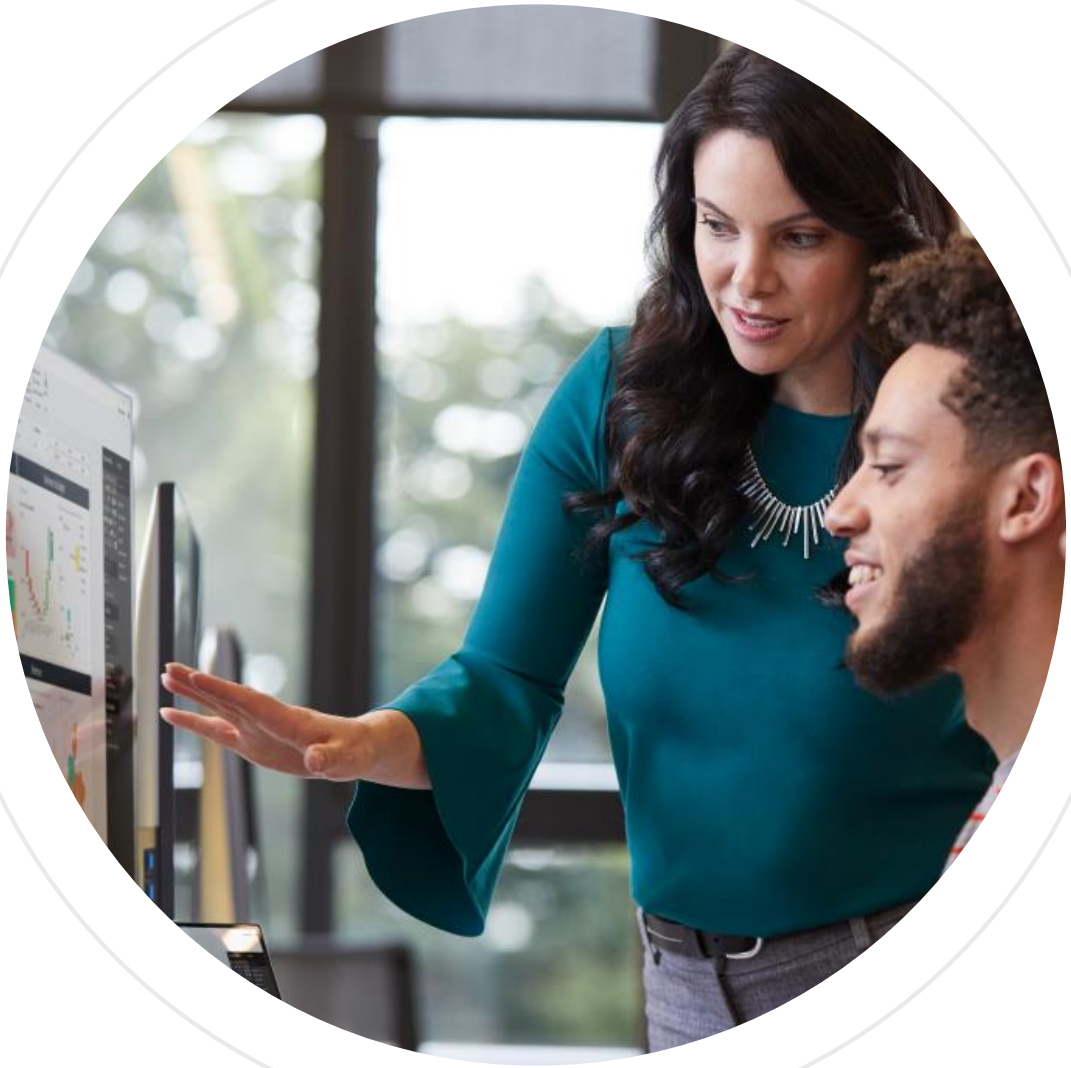
Six Sigma

with Minitab



Presenter

/arifmazumder

Mohammed Arif, PhD
Lead Data Scientist
Big Data | Machine Learning | AI

Mohammed Arif has more than fifteen (15) years of working experience in Information Communication and Technology (ICT) industry. The highlights of his career are more than six (7) years of holding various senior management and/or C-Level and had five (5) years of international ICT consultancy exposure in various countries (APAC and Australia), specially on Big Data, Data Engineering, Machine Learning and AI arena.

He is also Certified Trainer for Microsoft.

Introduction to Six Sigma Quality

Six Sigma was developed and used by Motorola in 1986, followed by General Electric and other companies in both manufacturing and service industries. Essentially, Six Sigma is a methodology that is used to improve an existing process (manufacturing/service).

In this section we will illustrates a number of terms used in this methodology and also gives an introduction to the **DMAIC (define-measure-analyze-improve-control)** approach used to improve a process.

Definitions

Definitions

Normal Distribution: If the histogram for a population of data looks like a symmetrical bell curve, it is likely that the data are normally distributed.

Process: A process is a set of tasks that convert inputs to outputs. For example, the process that manufactures the Toyota Camry (output) on an assembly line uses inputs such as capital, workforce, machines, facilities, and so on. If the output is a tangible product (e.g., a book by CRC Press), it is called a manufacturing process.

Process Mean: It is the average of the characteristic values of a population (e.g., an entire batch) of products produced by a process. For example, the process mean for the process that bottles 2-liter soda bottles may be 1.999 liters. The closer the process mean is to the process target, the better the process is.

Process Standard Deviation: It measures the average deviation of the characteristic values of a population of products from the process mean. For example, the process standard deviation for the process that bottles 2-liter soda bottles may be 0.002 liters. The lower the process standard deviation, the better (i.e., more consistent) the process is.

Definitions

Process Target: It is the best possible value for a product characteristic. For example, the process target for a process that bottles 2-liter soda bottles is 2 liters.

Process Tolerance: It is the deviation of the product characteristic value from the process target that the customer is willing to tolerate. For example, the process tolerance for the process that bottles 2-liter soda bottles may be ±0.003 liter. This means, the lower specification limit (LSL) is 1.997 liters (i.e., 2 – 0.003) and the upper specification limit (LSL) is 2.003 liters (i.e., 2 + 0.003). In other words, the customer in this example is satisfied if the volume is between 1.997 liters and 2.003 liters.

Quality: Quality is the extent to which customers are satisfied. If most of the customers are satisfied with a product, we can say that quality is high for the product as well as for the process that produces that product.

Voice of the Customer (VOC): This term is used to define the customer's wants or needs. It includes process target and process tolerance.

Voice of the Process (VOP): This term is used to define the process performance. It includes process mean and process standard deviation.

DMAIC Approach

DMAIC Approach

Six Sigma uses a systematic approach called DMAIC (define-measure-analyze-improve-control) to improve a given process.

Define: The problem is defined in this phase. For example, if the customers of a pizza delivery process are complaining about late delivery, then the problem may be defined as, “The average delivery time is long, and many of the customers are dissatisfied.”

Measure: In this phase, the current performance of the process is measured. For example, after taking a sample of deliveries, the average delivery time is calculated as 41 minutes, which is 11 minutes over the process target of 30 minutes.

Analyze: This phase analyzes the process to identify the root causes of the problem. For example, the pizza delivery process is analyzed to identify the potential causes (and then the root causes) of the long average delivery time.

Improve: In this phase, recommendations are made to minimize or eliminate the root causes of the problem, and then those recommendations are implemented to improve the process.

Control: This phase ensures that the improved process is controlled so that the process does not slide back to the previous problem.

Quality Analysis and Improvement Tools

Quality Analysis and Improvement Tools/Techniques

This section gives a brief definition/description of each of the quality analysis and improvement tools/techniques that will be used in the case studies.

Confidence Interval Estimation:

Confidence interval estimation is a technique to **estimate a population parameter** (such as population proportion) using sample data. The estimate is calculated for a given confidence level and is expressed as an interval. The **higher the confidence level** is, the **less precise the interval estimate**.

Hypothesis Testing

Hypothesis testing is a technique to **test whether there is enough statistical evidence to reject a claim**. Typically, the claim is expressed as the "null hypothesis," and an "alternative hypothesis" is considered to verify which of these two hypotheses is true. These two hypotheses are mutually exclusive (if one is true, the other one is not) and collectively exhaustive (no other hypothesis is possible).

Quality Analysis and Improvement Tools/Techniques

Chi-Square Analysis

Chi-square analysis is a type of hypothesis testing where a **sample statistic (called chi-square value) used in the test** is assumed to follow a chi-square distribution.

Process Capability Analysis

If USL is the upper specification limit for a process, LSL is the lower specification limit for a process, μ is the process mean, and σ is the process standard deviation, the following process capability ratios can measure process performance:

First-generation process capability ratio, $C_p = \frac{USL - LSL}{6\sigma}$

Second-generation process capability ratio with respect to LSL, $C_{pl} = \frac{\mu - LSL}{3\sigma}$

Second-generation process capability ratio with respect to USL, $C_{pu} = \frac{USL - \mu}{3\sigma}$

Second-generation process capability ratio, $C_{pk} = \text{MINIMUM of } \{C_{pl}, C_{pu}\}$

The higher the Cp and Cpk values are, the better the process is.

Quality Analysis and Improvement Tools/Techniques

Binary Logistic Regression

Binary logistic regression is a technique used to **predict the outcome of a binary categorical variable** with exactly two possible outcomes (e.g., Yes or No for whether a product is defective).

Item Analysis

Item analysis is used to check whether there is a **correlation** among categorical responses to multiple questions in a customer survey.

Cluster Analysis

Cluster analysis helps **group** customers into various clusters, using coordinate systems and Euclidean distances.

Mixture Design and Analysis of Experiments

Mixture design and analysis of experiments is a technique used to **optimize the proportion of each of the components** of a mixture such as a fuel mixture or a juice blend.

Quality Analysis and Improvement Tools/Techniques

Pareto Chart

The Pareto chart is a tool that can **prioritize problems using their frequencies**. Essentially, it is used to identify what few problems are causing the largest negative impact.

Cause-and-Effect Diagram

The cause-and-effect diagram is a tool that can **help identify root causes of a problem** (“effect”). It can also be used to divide numerous causes of a problem into several categories for easier analysis.

Gage Repeatability and Reproducibility Analysis

Gage repeatability and reproducibility (R&R) analysis is a technique that can verify whether the **measurement or inspection system used to collect the data** (e.g., number of defectives in a sample) is efficient. Repeatability measures the consistency of an inspector with himself or herself. Reproducibility measures the consistency of an inspector with other inspectors.

Quality Analysis and Improvement Tools/Techniques

Taguchi Design and Analysis of Experiments

Taguchi design and analysis of experiments is a technique that uses **signal-to-noise ratios** to check whether the variable for improvement depends on a certain set of controllable factors amid a set of uncontrollable (“noise”) factors.

Statistical Control Charts

It is important to **verify whether a process is stable and predictable** before measuring the performance of the process. Statistical control charts help in such verification.

Normality Test

The normality test is a tool to verify whether the collected data are **normally distributed**. This verification is important because many of the tools (e.g., statistical control charts) assume that the collected data are normally distributed.

Analysis of Variance (ANOVA)

Analysis of variance (ANOVA) is a technique that is often used in conjunction with **factorial design and analysis of experiments** to compare multiple population means while minimizing the error of rejecting a null hypothesis when the null hypothesis is true.