



DATA ANALYSIS TECHNIQUES FOR ENGINEERS & TECHNOLOGISTS



Introduction:

Corporate ethos which sees change as a survival necessity coupled with continual demands to achieve greater production efficiencies and reduced operating / maintenance costs, means that engineers and technologists are faced with ever-increasing plant and process performance targets.

As a consequence, more and more reliance is being placed upon the accurate and reliable analysis, representation and interpretation of data.

This course aims to provide engineers and technologists with the understanding and practical capabilities needed to convert data into information, and then to represent this information in ways that it can be readily exploited.

Training Methodology:

The course adopts an applications-oriented approach, minimizing the time spent on the mathematics of analysis and maximizing the time spent on the use of practical methods. Most of the time, delegates will work on computers, using

spreadsheets and the specialized analytical software MATLAB to investigate real data from a wide range of example applications. Delegates are encouraged to bring their own data with them on CD, and time will be set aside to address delegates' individual issues and concerns.

Who Should Attend?

The course has been designed for engineers, technologists, and professionals whose jobs involve the manipulation, representation, and analysis of data. Basic familiarity with PC's and in particular with Microsoft. Excel. is assumed. Aside from this, no other software familiarity or previous data analysis experience is required.

Course Objectives:

This Program aims to provide engineers and technologists with the understanding and practical capabilities needed to convert data into information, and then to represent this information in ways that it can be readily exploited.

- To provide delegates with a working vocabulary of analytical terms to enable them to converse with people who are experts in the areas of data analysis, statistics and probability, and to be able to read and comprehend common textbooks and journal articles in this field.
- To provide delegates with both an understanding and practical experience of a range of the more common analytical techniques and data representation methods, which have direct relevance to a wide range of analytical problems.
- To give delegates the ability to recognize which types of analysis are best suited to particular types of problems.
- To give delegates sufficient background and theoretical knowledge to be able to judge when an applied technique will likely lead to incorrect conclusions.
- To provide delegates with an overview of the main data analysis applications within engineering systems.

Course Outline:

The Basics

Sources of data, data sampling, data accuracy, data completeness, simple representations, dealing with practical issues. Workshop using centrifugal pump performance data.

Fundamental Statistics

Mean, average, median, mode, rank, variance, covariance, standard deviation, "lies, more lies and statistics", compensations for small sample sizes, descriptive statistics. Workshop using production data from a batch fermentation process.

Data Mining and Representation

Single, two and multi-dimensional data visualization, trend analysis, how to decide what it is that you want to see, box and whisker charts, common pitfalls and problems. Workshop using petrochemical plant control data.

Probability and Confidence

Probability theory, properties of distributions, expected values, setting confidence limits, risk and uncertainty, normal distribution, weibull distribution, binomial distribution, exponential distribution. Workshop using statistical process control data in the machinery protection system of a turbine-compressor installation.

Histograms & Frequency of Occurrence

Histograms, Pareto analysis (sorted histogram), cumulative percentage analysis, the law of diminishing return, percentile analysis. Workshop using historical failure data from a group of reciprocating compressors.

Frequency Analysis

The Fourier transform, periodic and a-periodic data, inverse transformation, practical implications of sample rate, dynamic range and amplitude resolution. Workshop using vibration data from a large gearbox.

Regression Analysis and Curve Fitting

Linear and non-linear regression, order; best fit; minimum variance, maximum likelihood, least squares fits, curve fitting theory, linear, exponential and polynomial curve fits, predictive methods. Workshop using failure data from large three phase induction motors, and remaining lifetime prediction.

Data Comparison

Correlation analysis, the autocorrelation function, Mahalanobis' distance, practical considerations of data set dimensionality. Workshop using diesel engine performance and pollutant emission data.

The Power of Excel and MATLAB

Pivot tables, the analytical toolbox, internet-based analysis tools and macros, dynamic spreadsheets, sensitivity analysis, visualization. Workshop involving step-by-step examples of the advanced capabilities of spreadsheets and the exploitation of ready-written resources.

Quality Control Applications

Terminology, control charts, statistical control, estimating the process mean and variation, capability indexes, control charts for attribute data. Workshop on constructing the \bar{x} and R charts for a milling process.

Reliability Evaluation Applications

Terminology, reliability definition and concepts, reliability functions, a reliability evaluation process. Workshop on evaluating the hazard rate, survivor function, failure density function, and cumulative failure distribution function for typical industrial equipment.