The importance of Design of Experiments (DOE)

Experimental design is an extensive subject by itself for discussion. It is a branch of applied statistics that deals with planning, conducting, analyzing and interpreting controlled tests to evaluate the factors that control the value of a parameter or group of parameters. It can be applied whenever you want to investigate a phenomenon in order to gain understanding or improve performance. For example, design of experiments (DOE) is used in the development and optimization of:

- manufacturing processes, and products, such as engines, semi-conductors
- analytical methods or instruments to determine their best or most sensitive performance
- synthesis of drug compounds in the pharmaceutical industry or in chemical research
- agricultural crop yields by studying various factors affecting the outcome, such as soil type, fertilizer, plant spacing, ambient temperatures, etc

Without a proper design of experiment (DOE), we may get stuck with the slow and tedious trial and error process which is time consuming and expensive. Hence, DOE is a powerful, structured and organized tool to define the relationship between factors (X) that affect a process or a chemical analysis, measured by its responses such as the yields of the process or the experimental means (Y). It involves designing a set of experiments, in which all relevant factors are varied systematically. The results of these experiments are then analysed to help identify optimal conditions, the factors that most influence the results, and those that do not. The existence of interaction and synergies between factors can also be determined.

It is interesting to know about the history of DOE. The concept was first conceived and developed by Sir Ronald A. Fischer in around 1930's. He was a brilliant mathematician and geneticist working on crop improvement in England. He designed and supervised field trials comparing fertilizers and seed varieties, amongst other things. Fischer encountered two enormous obstacles in his works, namely uncontrollable variation in soil type from plot to plot, and a limited number of plots available for any given trial. He finally solved these problems by the arrangement of the fertilizers or seed varieties in the field, and developed the renowned Analysis of Variance (ANOVA) approach. Fisher had demonstrated how taking the time to seriously consider the design and execution of an experiment before trying it helped avoiding frequently encountered problems in data analysis