Visualizing Data using R and ggplot2

R is an open source software for statistical analysis, machine learning and visualization. Thanks to a huge international community of developers, R is more powerful than many proprietary software packages.

ggplot2 has been developed by Hadley Wickham, arguably the best-known R developer, author of several books about R, and Chief Data Scientist at RStudio. ggplot2 is based on the grammar of graphics: diagrams are built step by step in layers and stored as objects. That way, they can be extended, manipulated and re-drawn in an automated way as desired. This makes visualizing data programmable, flexible, and powerful.

Topics

- Introduction to the 'Grammar of Graphics'
- Three basic layers: data, aesthetics, geometries
- Quick plots using simple syntax: qplot
- Introduction to the full ggplot2 syntax with more chart examples
- Tidy data, the underlying data structure
- Diagrams using facets (i.e. creating a collection of similar plots for subgroups with one plotting call)
- Statistical transformations, e.g. trend lines
- Outlook: themes, user-defined customizations, interactive diagrams, animations
- Overview of recommended extension packages

Prerequisites for taking this course

Please note that this is not a general introduction to R. A basic understanding of R and its data structures (data frames, vectors) is required, while prior knowledge of ggplot2 is not. Participants should be prepared to write R code.

Please note

This workshop cannot be credited as qualification measures at the Department of Mechanical Engineering. However you are welcome to participate.

This course will be held in English.