

Hadley Wickham

Data Visualisation in R: Harnessing the power of ggplot2 to produce elegant data graphics

London: 1st - 2nd November 2010

Mango Solutions is delighted to offer a one-off 2 day training course with Dr. Hadley Wickham, R Project Data Visualisation Guru and creator of ggplot 2. The course is a must for any R user looking to improve their data visualisation skills. Suitable for any level of R experience, it will especially benefit those who know how to get data in and out of R and who can do some basic modelling. Places will be limited on what will be a very popular course.

Course Content

Day One

Basic graphics

- How to create scatterplots, and how to add extra variables with aesthetics (like colour, shape and size) or facetting. Jittering, histograms and reordering. Coding strategy best practices.
- Data: fuel economy of US cars.

Graphics for large data

- Histograms and bar charts for displaying distributional summaries. More boxplots. Other techniques for overcoming overplotting when drawing scatterplots of large datasets.
- Data: prices and characteristics of 50,000 diamonds.

Data manipulation and transformation

- Group-wise summaries and transformations to add extra information to your plots. How to visualise time series.
- Data: trends in US baby names over the last 120 years.

Data analysis case study

- Iteration between graphics, transformation and modelling. More practice using dplyr, and combining it with other R functions.
- Data: US baby names

Day Two

Visualising space

- What is map data and how can you draw it? Working with shape files. Basic ggplot2 theory. How to combine multiple layers. Combining maps with data. Choropleths/thematic maps, proportional symbol maps.
- Data: Texas mortality and worldwide TB infections

ggplot2 theory and graphical critique

- How to critique a graphic (how is a graphic like pumpkin pie?) What is the layered grammar of graphics and why is it important? Using the grammar to create richer plots.

Polishing your plots for presentation

- Tweaking your plots for maximum presentation impact. Introduction to colour theory. Labels, legends and axes. Tweaking the plot themes.

Where to next

- General development advice.

For more information please contact: Training@mango-solutions.com
01249 767700 or visit www.mango-solutions.com

Venue to be confirmed but will be a central London location. Price for 2 day course £1,000 ex. VAT



Hadley Wickham is the Dobelman Family Junior Chair in Statistics at Rice University. He is an active member of the R community, has written and contributed to over 20 R packages and won the John Chambers Award for Statistical Computing for his work developing tools for data reshaping and visualisation. His research focuses on how to make data analysis better, faster and easier with a particular emphasis on the use of visualisation to better understand data and models.