Data Analysis Workshop

Description

The *Data Analysis* workshop enables laboratory-based life scientists to use the *R statistical programming environment* to analyse their own data.

R is an open-source cross-platform software tool that combines data manipulation, statistical modelling and visualisation. This workshop will focus on data manipulation and biostatistics modelling using relevant examples from the life sciences.

Using plenty of hands-on exercises, participants will learn about different data structures and functions in *R*, how to manage and ask specific questions of their data, and use the results of statistical tests. Programming capabilities that make *R* well-suited to data manipulation (e.g. `apply` functions) will also be introduced.

The workshop is available for two or three days. The three-day *Data Analysis* workshop allows for advanced topics to be explored in more detail. This includes advanced data management packages and concepts (e.g. `plyr` and `reshape2`) as well as automating common work-flows via scripting and an introduciton to control structures (e.g. looping and conditional statements). The three-day workshop also affords more time for students to work on their own data, with the goal of developing data analysis solutions as part of the workshop.

For further details, visit our web-site at `www.science-craft.com`.

The workshop instructor, Dr. Rick Scavetta, can be contacted directly via email at *rick.scavetta@science-craft.com*

Basic visualisations will be covered, but will be treated in more depth in the separate *Data Visualisation* workshop.

Requirements

The workshop takes a practical, hands-on approach to learning data analysis, but does not set-out to teach biostatistics. The implementation of basic descriptive statistics, in addition to estimation and inference methods and linear models, will be covered. Participants should be comfortable with computing and be familiar with basic biostatistics to take full advantage of the workshop.

> Participants are strongly encouraged to bring in their own data sets and computers for practical work.

Software

Students bringing their own computers should have the following cross-platform software pre-installed.

R – v3.0 or later (`http://www.r-project.org/`)

RStudio – v0.97 or later (`http://rstudio.org/download/desktop`)