

Data Handling in R: Practical Exercise

Version 3: August, 2011

Authored by Dr. Duncan Smallman on behalf of EDINA and Data Library, University of Edinburgh as part of the Research Data MANTRA (Management Training) Project, funded by JISC MRD Programme (2010-11): www.ed.ac.uk/is/data-library-projects/mantra

This practical consists of three exercises. Each exercise is provided as a separate document. The practicals are designed to be completed in order starting with practical 1 and finishing with practical 3. Please read the introductory material provided in this document and then start Exercise 1.

This work is licensed under the Creative Commons Attribution 2.5 UK: Scotland License.
To view a copy of this license, visit

<http://creativecommons.org/licenses/by/2.5/scotland/>



Introduction

R is a powerful graphics and statistical analysis package that is free to download from www.r-project.org. This web page is also a good place if you ever get stuck and need some help. R is purely a syntax package with limited menu based operations. This can seem daunting at first but does mean you have to know what you want to do and understand what you are doing.

This practical guide aims to provide you with an introduction and initial steps in managing and handling your data within R. The three exercises will cover how to:

1. input data into R
2. configure your data correctly to allow you to fully analyse them
3. work with your data and record your steps.

Requirements:

This practical assumes that you have a copy of R and a copy of MS Excel (or equivalent spreadsheet software). R as open source software can be freely obtained from www.r-project.org. Note that different installations vary and may have different libraries available and even different versions of R Commander.

Note of caution:

There will be some functions that have a summary statistics function. This practical will not run through all the various functions that can be used to obtain summary statistics of data.

Conventions:

Commands that are to be entered into R will appear in Courier New font just as it does in R.

Data sets:

All of the datasets required to complete the practical can be found in the compressed file on the MANTRA course website, called MANTRA-R-Practical-datafiles.zip. An R script is also included with the commands used in the practical.