# An Introduction to R

Di Cook
Iowa State University

R is an environment for data analysis. It contains the latest in statistical research methods, contributed by experts across the globe, and has the capacity to produce quality graphics. R is also complete programming language, so that new methods can be readily developed. This course is designed to help data analysts get started using R, focusing on basics and how R can be used in the process of data analysis.

This introductory course will be an elementary course for new users and practitioners having little experience with R. The focus will be on using R for exploring and modeling data. It will cover:

- An introduction to the R language (using graphics)
- The basics of handling data
- Basic math and vectors
- Producing data plots
- Simulation
- Reshaping data
- Being lazy with R
- Interactive graphics
- Statistical modeling of data
- Exploring multivariate data

We will practice all of the material discussed, so it is recommended that participants bring their laptops equipped with:

```
R version: 2.7.1 or later
Packages: MASS, ggplot2, rpart, randomForest, som, kohonen, nnet,
  e1071, mclust, norm, Hmisc, iplots, and dependencies of these packages.
More advanced participants might also like to have the ggobi
  software (2.1.7 or later) installed and rggobi 2.1.9 R package.
```

This book is recommended reading for the course:

> Modern Applied Statistics with S (4th ed), (2002)
> W.N. Venables and B. D. Ripley
> ISBN 0-387-95457-0


and this book might be interesting for advanced data analysts:

> Interactive and Dynamic Graphics for Data Analysis, 2007
> D. Cook and D. F. Swayne
> ISBN 0-387-71761-6