



DTC - STATISTICAL ANALYSIS USING R

Statistical Analysis using R: Level 1

OBJECTIVES

The Deerwalk Training Center- R Level 1- course is an integrated suite of software facilities for data manipulation, calculation and graphical display. It includes:

- An effective data handling and storage facility.
- A suite of operators for calculations on arrays, in particular matrices.
- A large, coherent, integrated collection of intermediate tools for data analysis.
- graphical facilities for data analysis and display either on-screen or on hardcopy, and
- A well-developed, simple and effective programming language which includes conditionals, loops, user-defined recursive functions and input and output facilities.

TARGET GROUP

Prerequisites:

- There are no hard prerequisites as such for learning R.
- What you need to understand is that R is one of the most popular analytics tool. But apart from being used for analytics,
- R is also a programming language.
- If you wish to use R for various analytical operations then some fundamental understanding of statistics would be really helpful.
- However, as a programming language, you don't need to know any of the other programming languages to start with R.

TRAINING METHOD

- The course is spread over 30 hours that consists of lecture and lab work. There will be approximately 10 hours of lectures and 20 hours of hands-on lab work.
- Lab exercises are mandatory, have a fixed deadline, and are graded. The course puts heavy emphasis on lab exercises because software programming can only be learnt well by explicitly putting into practice the principles that have been taught (i.e. in simpler terms – by doing lots and lots of coding). Late submission (past the deadline) of exercises incur some penalty from total points.
- Instructors may provide relevant lecture/lab notes to students as (and when) necessary in the form of printed handouts and or via emails.
- Instructors may provide supplementary code snippets to students via email or in lab class to support the theory and or lab material that is being taught.
- At the end of the course, students may have to give an exam (which will be optional), that will test their knowledge on the material covered during the course. This exam may be practical and/or theoretical and is mandatory for any student wishing to join a higher level.
- Students are graded on the basis of attendance, lab exercises and exam in the increasing order of importance.

COURSE DURATION

- 30 hours
- Classes
 - ✓ Morning/Evening

COURSE BREAKDOWN

Theory:

1. R Analytics
 - Introduction and preliminaries
 - The R environment
 - Related software and documentation
 - R and statistics
 - Using R interactively
 - An introductory session
 - Getting help with functions and features
 - R commands, case sensitivity, etc.
 - Executing commands from or diverting output to a file
 - Data permanency and removing objects
2. Simple manipulations; numbers and vector
 - Vectors and assignment
 - Vector arithmetic
 - Generating regular sequences
 - Logical vectors
 - Missing values
 - Character vectors
 - Index vectors; selecting and modifying subsets of a data set
 - Other types of objects
3. OBJECTS, THEIR MODES AND ATTRIBUTES
 - Intrinsic attributes: mode and length
 - Changing the length of an object
 - Getting and setting attributes
 - The class of an object
4. Ordered and unordered factors
 - A specific example
 - The function tapply() and ragged arrays
 - Ordered factors

5. ARRAYS AND MATRICES

- Arrays
- Array indexing. Subsections of an array
- Index matrices
- The array() function
- Mixed vector and array arithmetic. The recycling rule
- The outer product of two arrays
- Generalized transpose of an array
- Matrix facilities
- Matrix
- Linear equations and inversion
- Forming partitioned matrices, cbind() and rbind5.9 The concatenation function, c(), with arrays
- Frequency tables from factors

6. LISTS AND DATA FRAMES

- Lists
- Constructing and modifying lists
- Concatenating lists

7. DATA frames

- Making data frames
- attach() and detach()
- Working with data frames
- Attaching arbitrary lists
- Managing the search path

8. READING DATA FROM FILES

- The read.table() function
- The scan() function
- Accessing builtin datasets
- Loading data from other R packages

9. EDITING DATA

10. GROUPING, LOOPS AND CONDITIONAL EXECUTION

- Grouped expressions
- Control statements
- Conditional execution: if statements
- Repetitive execution: for loops, repeat and while

11. WRITING YOUR OWN FUNCTIONS

- Simple examples
- Defining new binary operators
- Named arguments and defaults
- The '...' argument
- Assignments within functions
- Scope

- Classes, generic functions and object orientation
12. Graphical procedures
 13. High-level plotting commands
 - The plot() function
 - Displaying multivariate data
 - Display graphics
 - Arguments to high-level plotting functions
 - Low-level plotting commands
 - Mathematical annotation
 - Hershey vector fonts
 - Interacting with graphics
 - Using graphics parameters
 - Permanent changes: The par() function
 - Temporary changes: Arguments to graphics functions
 14. Packages
 - Standard packages
 - Contributed packages and CRAN
 - Namespaces

LABS

Lab assignments will focus on the practice and mastery of contents covered in the lectures; and introduce critical and fundamental problem-solving techniques to the students.

DISCLAIMER

Please note that Deerwalk Training Center reserves the right to change the course syllabus of DTC - R course – Level 1 course at any time without prior notification.

Statistical Analysis using R: Level 2

OBJECTIVES

The Deerwalk Training Center- R Level 2- course is an integrated suite of software facilities for data manipulation, calculation and graphical display. It includes:

- An effective data handling and storage facility.
- A suite of operators for calculations on arrays, in particular matrices.
- A large, coherent, integrated collection of intermediate tools for data analysis.
- graphical facilities for data analysis and display either on-screen or on hardcopy, and
- A well-developed, simple and effective programming language which includes conditionals, loops, user-defined recursive functions and input and output facilities.

TARGET GROUP

Prerequisites:

- There are no hard prerequisites as such for learning R.
- What you need to understand is that R is one of the most popular analytics tool. But apart from being used for analytics,
- R is also a programming language.
- If you wish to use R for various analytical operations then some fundamental understanding of statistics would be really helpful.
- However, as a programming language, you don't need to know any of the other programming languages to start with R.

TRAINING METHOD

- The course is spread over 30 hours that consists of approximately 10 hours of lecture and 20 hours of hands-on lab work.
- Lab exercises are mandatory, have a fixed deadline, and are graded. The course puts heavy emphasis on lab exercises because software programming can only be learnt well by explicitly putting into practice the principles that have been taught (i.e. in simpler terms – by doing lots and lots of coding). Late submission (past the deadline) of exercises incur some penalty from total points.
- Instructors may provide relevant lecture/lab notes to students as (and when) necessary in the form of printed handouts and or via emails.
- Instructors may provide supplementary code snippets to students via email or in lab class to support the theory and or lab material that is being taught.
- At the end of the course, students may have to give an exam (which will be optional), that will test their knowledge on the material covered during the course. This exam may be practical and/or theoretical and is mandatory for any student wishing to join a higher level.
- Students are graded on the basis of attendance, lab exercises and exam in the increasing order of importance.

COURSE DURATION

- 30 hours
- Classes
 - ✓ Morning/Evening

COURSE BREAKDOWN

1. DATA TRANSFORMATION USING DPLYR

- Summarize Cases
- Group Cases
- Manipulate Cases
- Extract Cases
- Arrange Cases
- Manipulate Variables
- Extract Variables
- Make New Variables
- Vectorized Functions
- Summary Functions
- Combine Variables
- Combine Cases

2. R Markdown

- Knitr
- .Rmd files
- Interactive Documents
- Parameters
- Pandoc's Markdown
- YAML
- Re-using Template
- Table Suggestions

3. BUILDING APPLICATION

- Part 1 - How to build a Shiny app
- Introduction
- R
- App architecture
- App template
- Inputs and outputs
- The server function
- Sharing apps
- Shinyapps.io
- Shiny servers

4. Part 2 - How to customize reactions

- Introduction
- Review of Part 1

- Reactivity
 - Reactive values
 - Reactive functions
 - render*()
 - reactive()
 - isolate()
 - observeEvent()
 - eventReactive()
 - reactiveValues()
 - Parting tips
5. PART 3 - HOW TO CUSTOMIZE APPEARANCE
- Introduction
 - Review of Parts 1 and 2
 - HTML UI
 - Adding static content
 - Building layouts
 - Panels and tabsets
 - Prepackaged layouts
 - CSS
6. R AS A DATABASE MANAGEMENT SYSTEM (DBMS)
- Data Definition Language (DDL)
 - Data Manipulation Language (DML)
 - Writing Functions
 - Cursor and Views
 - Big Data in R
 - Automation using R

LABS

Lab assignments will focus on the practice and mastery of contents covered in the lectures; and introduce critical and fundamental problem-solving techniques to the students.

DISCLAIMER

Please note that Deerwalk Training Center reserves the right to change the course syllabus of DTC - R course – Level 2 course at any time without prior notification.