

# **INTRODUCTION TO STATISTICAL ANALYSIS USING R**

## **Statistical Analysis using R: Level 1**

Nature of the course: Theory + Practical

Total hours per day: 2 Hours

Course duration: 5 Weeks

### **Course Summary**

The R course at the Deerwalk Training Center offers a comprehensive set of software tools for data processing, calculation, and graphical display. It consists of the following components: an efficient data handling and storage facility, a set of operators for working with arrays, particularly matrices. This course covers a well-developed, simple, and effective programming language with conditionals, loops, user-defined recursive functions, and input and output facilities.

### **Completion Criteria**

After fulfilling all of the following criteria, the student will be deemed to have finished the Module:

- Has attended 90% of all classes held
- Has received an average grade of 80% on all assignments
- Has received an average of 60% in assessments
- The tutor believes the student has grasped all of the concepts and is ready to go on to the second module.

### **Required Textbooks**

- Andrie de Vries and Joris Meys, “R for Dummies”, Wiley.
- Jum Albert and Maria Rizzo, “R by Example”, Springer Media
- Michael J. Crawley, “The R Book”, Wiley.

### **Prerequisites**

- There are no specific prerequisites for learning R.
- If you want to utilize R for a variety of analytical tasks, you'll need to

have a basic understanding of statistics.

- However, to get started with R as a programming language, you don't need to know any of the other programming languages.

## **Course Details**

### **Week I**

#### **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

#### **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

#### **Objects, Modes And Attributes**

- Intrinsic attributes: mode and length
- Changing the length of an object
- Getting and setting attributes
- The class of an object

### **Week II**

#### **Ordered And Unordered Factors**

- A specific example

- The function `apply()` and ragged arrays
- Ordered factors

## **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
- Linear equations and inversion
- Forming partitioned matrices, `cbind()` and `rbind()`
- The concatenation function, `c()`, with arrays
- Frequency tables from factors

## **Lists And Data Frames**

- Lists
- Constructing and modifying lists
- Concatenating lists

## **Week III**

### **Data Frames**

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

### **Reading Data From Files**

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

## **Week IV**

### **Grouping, Loops And Conditional Execution**

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

### **Writing Your Own Functions**

- Assignments within functions
- Scope
- Classes, generic functions and object orientation
- Defining new binary operators
- Named arguments and defaults
- Simple examples

## **Week V**

### **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

### **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.

## **Intermediate Statistical Analysis using R: Level 2**

Nature of the Course: Theory + Practical

Total Hours per Day: 2 Hours

Course Duration: 3 Weeks

### **Course Summary**

The R course at the Deerwalk Training Center offers a comprehensive set of software tools for data processing, calculation, and graphical display. It consists of the following components: an efficient data handling and storage facility, a set of operators for working with arrays, particularly matrices. This course covers a well-developed, simple, and effective programming language with conditionals, loops, user-defined recursive functions, and input and output facilities.

### **Completion Criteria**

After fulfilling all of the following criteria, the student will be deemed to have finished the Module:

- Has attended 90% of all classes held
- Has received an average grade of 80% on all assignments
- Has received an average of 60% in assessments
- The tutor believes the student has grasped all of the concepts and is ready to go on to the second module.

### **Required Textbooks**

- Jum Albert and Maria Rizzo, “R by Example”, Springer Media
- Michael J. Crawley, “The R Book”, Wiley.

### **Prerequisites**

- There are no specific prerequisites for learning R.
- If you want to utilize R for a variety of analytical tasks, you'll need to have a basic understanding of statistics.
- However, to get started with R as a programming language, you don't need to know any of the other programming languages.

## Course Details

### Week I

#### 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

### R Markdown

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

### Week II

#### Building Application

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

- Shiny servers

## **Part 2 - How To Customize Reactions**

- Introduction
- Review of Part 1
- Reactivity
- Reactive values
- Reactive functions
- render\*()
- reactive ()
- isolate ()
- observe Event ()
- event Reactive ()
- reactive Values ()
- Parting tips

## **Week III**

### **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

### **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.

## **Learning Outcomes**

- Import, examine, manipulate, and summarize data sets in R
- Explore data sets to develop testable hypotheses and find applicable statistical tests
- Use R to do relevant statistical tests Create and edit visualizations with R
- Learn the fundamentals of R programming, including constructions, control statements, and string functions
- Identify the key terminologies, concepts, and techniques used in statistical analysis.
- Learn how to use R programming for text processing
- Able to understand and implement R programming from a statistical standpoint.