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:

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

Required Text Books

1. Andrie de Vries and Joris Meys, "R for Dummies", Wiley.
2. Jum Albert and Maria Rizzo, "R by Example", Springer Media
3. 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 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

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 2

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 rbind5.9 The concatenation function, c(), with arrays
- Frequency tables from factors

LISTS AND DATA FRAMES
- Lists
- Constructing and modifying lists
- Concatenating lists

WEEK 3
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 4
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 5
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:

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

Required Text Books

1. Jum Albert and Maria Rizzo, "R by Example", Springer Media
2. 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 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

R Markdown

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

WEEK 2

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 3
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.