



FOUNDATIONS OF R SOFTWARE

PROF. SHALABH

Department Of Mathematics And Statistics
IIT Kanpur

PRE-REQUISITES : Mathematics background up to class 12 is needed. Having some preliminary knowledge will be helpful but not necessarily mandatory.

INTENDED AUDIENCE : UG students of Science and Engineering. Students of humanities with basic mathematical background can also do it. Working professionals in analytics and anyone involved in programming, mathematical and statistical computations, simulations and data analysis can also do it.

INDUSTRY SUPPORT : All industries involved in mathematical and statistical computations, programming and simulations and having R & D set up will use this course.

COURSE OUTLINE :

Any scientific task without the knowledge of software is difficult to complete in the current scenario. R is free software that is capable of mathematical and statistical computing. It has its own programming language, built-in functions, and graphical capabilities to perform any specialized task in mathematics, statistics, and other areas. We intend to learn the fundamentals of R software in this course.

ABOUT INSTRUCTOR :

Prof. Shalabh is a Professor of Statistics and Data Science at IIT Kanpur. His research areas of interest are linear models, regression analysis, and econometrics. He has more than 25 years of experience in teaching and research. He has developed several web-based and MOOC courses in NPTEL and has conducted workshops on statistics for teachers, researchers, and practitioners. He has received several national and international awards and fellowships. He has authored more than 100 research papers in national and international journals. He has written four books, and one of the books on linear models is coauthored with Prof. C.R. Rao. Another seminal book on Statistics with R software has been downloaded more than 5.4 million times. He has completed several projects and has played an important role in propagating the knowledge of R software in the country.

COURSE PLAN :

Week 1: Why R and Installation Procedure; Help, Demonstration, and Examples in R; Packages and Libraries in R; Command Line and Data Editor

Week 2: Introduction to R Studio; Basic Operations in R; R as a Calculator with Scalars and Data Vectors : Addition, Subtraction, Multiplication & Division

Week 3: Calculations with Data Vectors :Addition, Subtraction, Multiplication & Division; R as a Calculator with Scalars and Data Vectors : Power operations, Integer and Modulo divisions; Built in Functions and Assignments; Matrices

Week 4: Matrix Operations – Row, Column & Other Operations; Matrix Operations – Access and Mathematical Operations; Matrix Operations – Mathematical and Other Operations; Logical Operators; Relational and Logical Operators

Week 5: Missing Data Handling, Conditional Executions – If and If-Else, Conditional Executions – Nested if else if and ifelse; Functions for Conditional Executions – switch and which commands

Week 6: Loops – for loop; Loops – while and repeat; Functions; Sequences

Week 7: More Sequences and Other Operations; Sequences of Dates and Alphabets; Repeats; Sorting, Ordering and Mode; Lists

Week 8: Operations on Lists; Vector Indexing; Factors; Factors – Class and Unclass; Strings – Display and Formatting : Print and Format Function

Week 9: Strings – Display and Formatting: Print and Format with Concatenate; Paste Function; String Splitting; Manipulations with Strings and Alphabets

Week 10: Strings – Display and Formatting: Substitution and Replacement of Strings; Data Frames; Data Frames: Creation and Operations; Data Frames: Combining and Merging

Week 11: Data Handling: Importing and Reading CSV and Tabular Data Files; Importing and Reading EXCEL and other Data Files; Saving and Writing Data Files; Introduction to Statistical Functions: Introduction, Frequencies and Partition Values;

Week 12: Graphics: Scatter Plot and Bar Plots; Sub-Divided Bar Plots and Pie Diagram , Histogram; Bivariate and Three Dimensional Scatter Plots, Some Examples of R Programming