



INTRODUCTION TO R SOFTWARE

PROF. SHALABH

Department of Mathematics
IIT Kanpur

TYPE OF COURSE : Rerun | Core | UG

COURSE DURATION : 8 weeks (26 Jul'21 - 17 Sep'21)

EXAM DATE : 26 Sep 2021

PREREQUISITES : 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.

INDUSTRIES APPLICABLE TO : 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 imagine and complete in the current scenario. R is a free software that is capable of handling mathematical and statistical manipulations. It has its own programming language as well as built in functions to perform any specialized task. We intend to learn the basics of R software in this course.

ABOUT INSTRUCTOR

Prof. Shalabh is a Professor of Statistics at IIT Kanpur. His research areas of interest are linear models, regression analysis and econometrics. He has more than 22 years of experience in teaching and research. He has developed several web based NPTEL courses including on regression analysis and has conducted several workshops on statistics for teachers, researchers and practitioners. He has received several national and international award and fellowships. He has authored more than 70 research papers in national and international journals. He has written four books and one of the book on linear models is coauthored with Prof. C.R. Rao.

COURSE PLAN

Week 1: Basic fundamentals, installation and use of software, data editing, use of R as a calculator, functions and assignments.

Week 2: Use of R as a calculator, functions and matrix operations, missing data and logical operators.

Week 3: conditional executions and loops, data management with sequences.

Week 4: Data management with repeats, sorting, ordering, and lists.

Week 5: Vector indexing, factors, Data management with strings, display and formatting.

Week 6: Data management with display paste, split, find and replacement, manipulations with alphabets, evaluation of strings, data frames.

Week 7: Data frames, import of external data in various file formats, statistical functions, compilation of data.

Week 8: Graphics and plots, statistical functions for central tendency, variation, skewness and kurtosis, handling of bivariate data through graphics, correlations, programming and illustration with examples.