



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

15.05.2023

Submitted to the Principal,

SUB: REPORT FOR WORKSHOP ON “Data Analytics using R-Programming” for 3 days-reg

“Data Analytics using R-Programming”, workshop in collaboration with ISTE Chapter for 3 days (11.05.2023-13.05.2023) was organised for III-year students who belong to the Department of Computer Science and Engineering with the venue as CSE laboratory. The session started with a welcome address by the Head of the Department, Dr. Smitha J A and the session was handed over to the guest, Dr. M. Senbagavalli, Associate Professor, Department of Computer Science and Information Technology, Alliance College of Engineering and Design, Bangalore.

R is a programming language for statistical computing and graphics supported by the R Core Team and the R Foundation for Statistical Computing. Created by statisticians Ross Ihaka and Robert Gentleman, R is used among data miners, bioinformaticians and statisticians for data analysis and developing statistical software. The core R language is augmented by a large number of extension packages containing reusable code and documentation. The guest gave in-depth knowledge to the students about R-programming introduction on the first session as theory.

The second day was given a theory session with data processing techniques of R-programming. R's data structures include vectors, arrays, lists, data frames and environments. Vectors are ordered collections of values and can be mapped to arrays of one or more dimensions in a column major order. That is, given an ordered collection of dimensions, one fills in values along the first dimension first, then fills in one-dimensional arrays across the second dimension, and so on. R supports array arithmetics and in this regard is like languages such as APL and MATLAB. The special case of an array with two dimensions is called a matrix. Lists serve as collections of objects that do not necessarily have the same data type. Data frames contain a list of vectors of the same length, plus a unique set of row names. R has no scalar data type. Instead, a scalar is represented as a length-one vector.

R and its libraries implement various statistical techniques, including linear, generalized linear and nonlinear modelling, classical statistical tests, spatial and time-series analysis, classification, clustering, and others. For computationally intensive tasks, C, C++, and Fortran code can be linked and called at run time. Another of R's strengths is static graphics; it can produce publication-quality graphs that include mathematical symbols. The hands-on session with data processing techniques is given to students at CSE lab with an open-source tool.

The third day was given a theory session regarding R-programming packages. R's capabilities are extended through user-created packages, which offer statistical techniques, graphical devices, import/export, reporting (RMarkdown, knitr, Sweave), etc. These packages and their easy installation and use has been cited as driving the language's widespread adoption in data science. The packaging system is also used by researchers to organize research data, code, and report files in a systematic way for sharing and archiving.

At the last session of the third day, the workshop ended with valediction addressed by Dr. Smitha J A, Head of the Department, CSE. The guest was given a letter of appreciation and amount of honorarium. The students were given a certificate of appreciation after attending their assessment. The students gained clear knowledge about R-programming and experience of hands-on sessions.



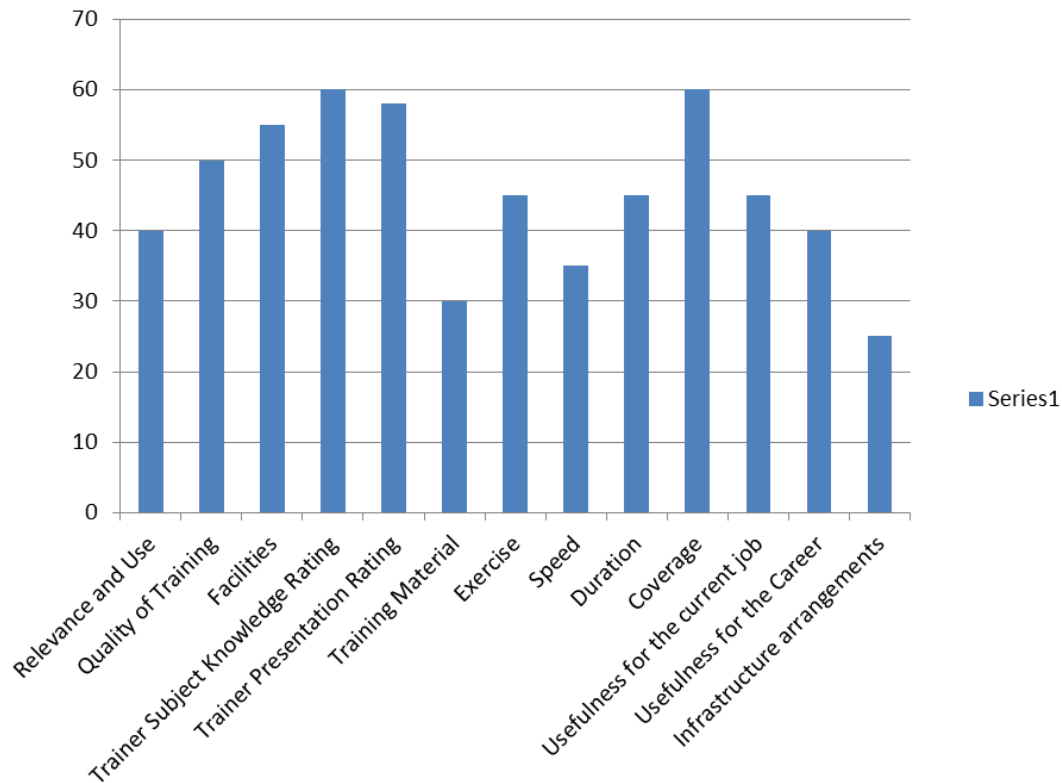
Sri
SAIRAM
COLLEGE OF ENGINEERING

Accredited by NAAC
ISO 9001:2015 Certified Institution
Approved by AICTE, New Delhi
Affiliated to Visvesvaraya Technological University
www.sairamce.edu.in





Feedback Analysis for Data Analytics Using R- Programming



Co-Ordinator

HoD