

Anekant Education Society's

**Tuljaram Chaturchand College
of
Arts , Science and Commerce**
(Autonomous)

Baramati

Department Of Computer Science

Ms. Kalyani Waman Londhe.
M.Sc. (Computer Science) NET



What is R ?

- 
- R is a programming language.
 - It is environment for statistical computing and graphics together.
 - It has Extensive collection of packages of data mining.
 - R was initially written by Robert Gentleman and Ross Ihaka, at the University of Auckland, New Zealand.
 - It is developed by the *R Development Core Team*.

- 
- R is named partly after the first names of the first two R authors (Robert Gentleman and Ross Ihaka)
 - Hence , it is also known as R&R.
 - R is an object-oriented programming language.
 - This means that everything what is done with R can be saved as an object.
 - Every object has a class.
 - It describes what the object contains and what each function does.

- 
- It is similar to S language.
 - It is different implementation of S language combined with lexical scoping semantics, inspired by Scheme.
 - There are some important differences but Most of the code written in S language can work unaltered under R.
 - R is suited to challenging tasks associated with data mining.
 - It is highly extensible through functions and extensions.

- 
- R probably has the biggest library of data mining modules.
 - These modules are more mature than Python's Scikit-learn.
 - R contains many specialized application packages and libraries for a huge number of statistical, mathematical and other methods.
 - All data is processed in memory, hence working with larger volumes of data is very slow.

- 
- The R language is widely used among statisticians and data miners for developing statistical software and data analysis.
 - Although R has a command line interface, there are several third-party graphical user interfaces, such as RStudio and Integrated Development Environment.
 - Application of R as a programming language and statistical software is much more than a supplement to Stata, SAS, and SPSS.

Features of R language

- It has an effective data handling and storage facility.
- It has a suite of operators for calculations on arrays in particular matrices.
- R is available as free software under the terms of free software foundations GNU General Public License in source code form.
- R can be used for linear and non-linear modeling, classical statistical tests, time-series analysis, classification and clustering.

- 
- It has a large coherent, integrated collection of intermediate tools for data analysis.
 - It has various Graphical facilities for data analysis and display either on screen or on hard copy.
 - It complies and runs on a variety of Unix platforms and similar systems including FreeBSD , Linux , windows and Mac'OS.
 - R is a simple and very powerful data mining and statistical data processing tool.
 - It provides users with an entirely new, rich and powerful tool applicable in almost every field of research.

RATTLE PACKAGE

- Rattle stands for R Analytical Tool To Learn Easily.
- Rattle is one of the package from R.
- It is graphical data mining application.
- Rattle uses the Gnome graphical user interface as provided through the RGtk2 package.
- It is written in and provided a pathway to R.

- 
- Understanding of R is not necessary to start with the Rattle.
 - It has been developed specifically to ease the transition from basic data mining, as necessarily offered by GUI's.
 - Rattle brings together a multitude of R packages.
 - These packages are very useful for data miners.
 - These are not very easy to use for novice.

- 
- Rattle is used for teaching data mining at numerous universities.
 - It is daily used by consultants and data mining teams world wide.
 - This is also available as a product within Information Builder's Web Focus business intelligence suite as Rstat.

- 
- Rattle is one of the several open source data mining tools.
 - Many of these tools are also directly available within R through packages like Rweka and arules.



Thank You