



# **SNS COLLEGE OF ENGINEERING**

Kurumbapalayam (Po), Coimbatore – 641 107

**An Autonomous Institution**

Accredited by NAAC – UGC with 'A' Grade

Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai



## **DEPARTMENT OF COMPUTER SCIENCE AND TECHNOLOGY**

**COURSE NAME :19CS407 DATA ANALYTICS WITH R**  
**II YEAR /IV SEMESTER**

**Unit 4- R PROGRAMMING BASICS**

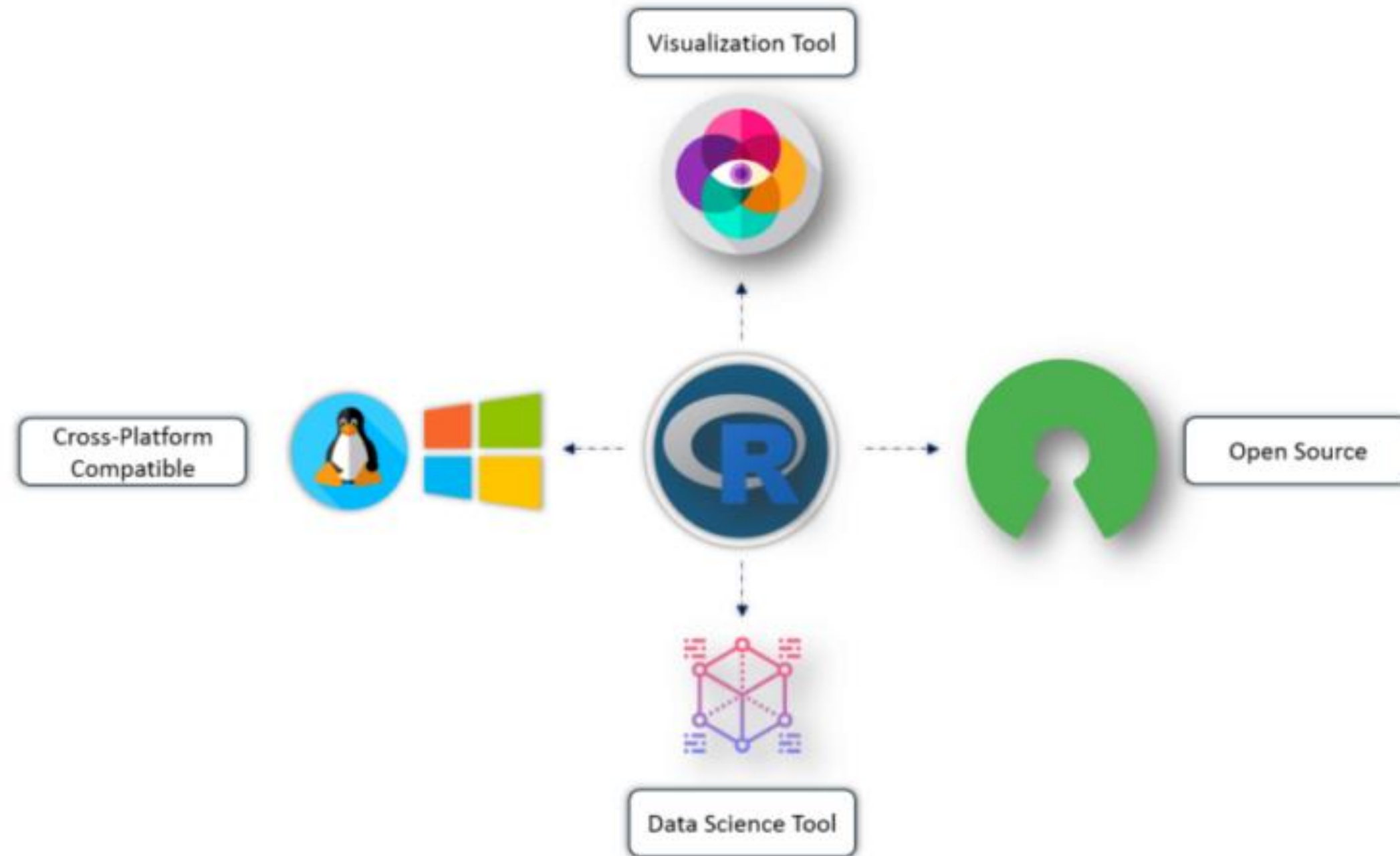
**Topic : Overview of R programming, Environment setup  
with R Studio**



# QR

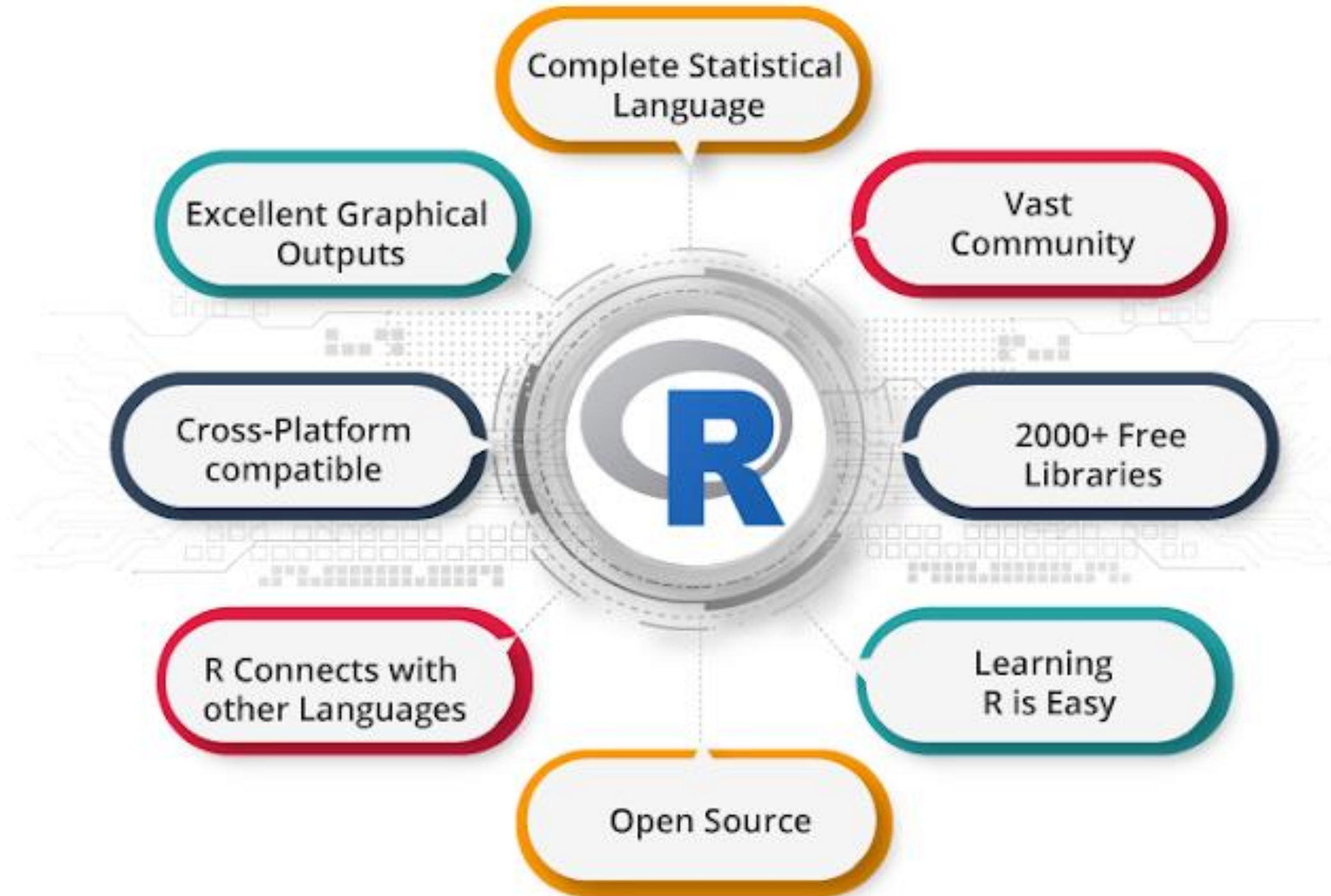


# Overview of R programming





# Overview of R programming





# Is R programming language?

- ✓ If you ever had a question like is an R programming language? then my answer for it should be a big YES.
- ✓ The R language is particularly made to suit statistical modeling, data analysis, and more computation heavy tasks.
- ✓ It's amazing features, tools and graphics made it even more compatible and scalable to solve real-world problems.





# Is R programming language?



- ✓ The R language is an open-source language available under GNU general public license.
- ✓ The key advantages of the R language are its massive developer community, amazing analytical tools, and the capability of performing analytical tasks over tons of data.
- ✓ You can use R in Windows, macOS, Linux, and Unix devices.



# R - Overview



- ✓ R is a programming language and software environment for statistical analysis, graphics representation and reporting. R was created by Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand, and is currently developed by the R Development Core Team.
- ✓ The core of R is an interpreted computer language which allows branching and looping as well as modular programming using functions. R allows integration with the procedures written in the C, C++, .Net, Python or FORTRAN languages for efficiency



# R - Overview



- ✓ R is freely available under the GNU General Public License, and pre-compiled binary versions are provided for various operating systems like Linux, Windows and Mac.
- ✓ R is free software distributed under a GNU-style copy left, and an official part of the GNU project called GNU S.





# Evolution of R



- ✓ R was initially written by Ross Ihaka and Robert Gentleman at the Department of Statistics of the University of Auckland in Auckland, New Zealand. R made its first appearance in 1993.
- ✓ A large group of individuals has contributed to R by sending code and bug reports.
- ✓ Since mid-1997 there has been a core group (the "R Core Team") who can modify the R source code archive.



# Features of R



- ✓ As stated earlier, R is a programming language and software environment for statistical analysis, graphics representation and reporting. The following are the important features of R –
- ✓ R is a well-developed, simple and effective programming language which includes conditionals, loops, user defined recursive functions and input and output facilities.
- ✓ R has an effective data handling and storage facility,
- ✓ R provides a suite of operators for calculations on arrays, lists, vectors and matrices.



# Features of R



- ✓ R provides a large, coherent and integrated collection of tools for data analysis.
- ✓ R provides graphical facilities for data analysis and display either directly at the computer or printing at the papers.
- ✓ As a conclusion, R is world's most widely used statistics programming language. It's the # 1 choice of data scientists and supported by a vibrant and talented community of contributors. R is taught in universities and deployed in mission critical business applications.



# Applications of R



- ✓ Facebook
- ✓ Google
- ✓ Twitter
- ✓ HRDAG
- ✓ Sunlight Foundation
- ✓ RealClimate
- ✓ NDAA
- ✓ XBOX ONE
- ✓ ANZ
- ✓ FDA



# Why use R Programming?

- ✓ The important task in data science is the way we deal with the data: clean, feature engineering, feature selection, and import. It should be our primary focus.
- ✓ Data scientist job is to understand the data, manipulate it, and expose the best approach. For machine learning, the best algorithms can be implemented with R.
- ✓ Keras and TensorFlow allow us to create high-end machine learning techniques. R has a package to perform Xgboost. Xgboost is one of the best algorithms for Kaggle competition.





# Why use R Programming?



- ✓ R communicate with the other languages and possibly calls Python, Java, C++. The big data world is also accessible to R. We can connect R with different databases like Spark or Hadoop.



# Features of R programming



- ✓ It is a simple and effective programming language which has been well developed.
- ✓ It is data analysis software.
- ✓ It is a well-designed, easy, and effective language which has the concepts of user-defined, looping, conditional, and various I/O facilities.
- ✓ It has a consistent and incorporated set of tools which are used for data analysis.



# Features of R programming



- ✓ For different types of calculation on arrays, lists and vectors, R contains a suite of operators.
- ✓ It provides effective data handling and storage facility.
- ✓ It is an open-source, powerful, and highly extensible software.
- ✓ It provides highly extensible graphical techniques.
- ✓ It allows us to perform multiple calculations using vectors.
- ✓ R is an interpreted language.



# Assessment 1





# References



1. João Moreira, Andre Carvalho, Tomás Horvath – “A General Introduction to Data Analytics” – Wiley -2018

**Thank You**