

**IK Gujral Punjab Technical University, Kapurthala**  
**B. Tech, Computer Science & Engineering, with AI & ML**

---

Course Code: BTES 501-20	Course Title: Statistical Computing Techniques using R	3L:0T:0P	3 Credits
--------------------------	---	----------	-----------

**Course Contents:**

**UNIT 1:** **CO1, CO2 [8 Hrs.]**

General introduction to computing, Using R as a calculator, Numbers, words and logicals; missing values (NA), Vectors and their attributes (names, length, type), System- and user-defined objects, Accessing data (data()). Data in the system and data outside the system (read.table, scan)

**UNIT 2:** **CO1, CO2 [10 Hrs.]**

First steps in graphics, The basics of R syntax, The R workspace, Matrices and lists, Subsetting System-defined functions; the help system, Errors and warnings; coherence of the workspace Data input and output; interface with other software packages, Writing your own code; R script Good programming practice, R syntax -- further steps The parentheses and brackets; =, == and <-

Apply-type functions Compiling and applying functions Documentation, Conditional statements Loops and iterations

**UNIT 3:** **CO1, CO2 [8 Hrs.]**

Exploratory data analysis, Range, summary, mean, variance, median, sd, histogram, box plot, scatterplot

Probability distributions, Random number generation Distributions, the practice of simulation.

**UNIT 4:** **CO1, CO2, CO3 [8 Hrs.]**

Statistical functions in R, Statistical inference, contingency tables, chi-square goodness of fit, regression, generalized linear models, advanced modelling methods, the bootstrap method to compute s.e.f

**UNIT 5:** **CO1, CO3 [8 Hrs.]**

Graphics; beyond the basics Graphics and tables, Working with larger datasets, Principles of exploratory data analysis (big data analysis)

Dataframes in R, Defining your own classes and operations Models and methods in R, Customising the user's environment

**Reference Books:**

1. Matloff, N. (2011). The Art of R Programming: A Tour of Statistical Software Design, William
2. Philip H. Pollock (2014). An R Companion to Political Analysis, CQ Press
3. Chihara, L. and Hesterberg, T. (2011), Mathematical statistics with resampling and R, Wiley
4. Lander, J. P. (2014) R for Everyone: Advanced Analytics and Graphics, Addison-Wesley Data & Analytics Series

### **Course Outcomes:**

At the end of the course, students will have learned:

CO1: To use a fundamental tool for computing in the practice of quantitative analytical methods (the ‘paper-and-pencil’ tool of the 21st century), that can work for the small jobs (like a pocket calculator) as well as for the big jobs (complex statistical data analysis).

CO2: Programming, data handling, transformations, subsetting, exploratory data analysis, probability distributions and simulations, regression and linear models, summarising data, how to handle large data sets, effective graphics.

CO3: Modern concepts of statistics based on simulations and writing a report of a quantitative analysis.

-----

Course Code: BTES 502-20	Course Title: Statistical Computing Techniques using R lab	0L:0T:2P	1 Credits
--------------------------	---	----------	-----------

### **Details of Experiments**

1. Familiarization of environments in R.
2. Perform simple arithmetic's using R.
3. Perform basic R functions.
4. Use various graphical techniques in EDA.
5. Create different charts for visualization of given set of data.
6. Find the mean, median, standard deviation and quartiles of a set of observations.
7. Find the Skewness and Kurtosis of a given dataset distribution.
8. Given the scenario, implement the Bayes rule by finding the posterior probability.
9. Find the mass function of a binomial distribution with  $n=20, p=0.4$ . Also draw the graphs of the mass function and cumulative distribution function.
10. Generate and draw the cdf and pdf of a normal distribution with mean=10 and standard deviation=3. Use values of  $x$  from 0 to 20 in intervals of 1.
11. Construct a scatter plot to investigate the relationship between two variables.
12. Perform the Z- test for single proportion, single mean etc.
13. Calculate the regression coefficient and obtain the lines of regression for the given data.
14. Compute confidence intervals for the mean when the standard deviation is known.
15. Perform F test
16. Perform Chi-Square test.