**Paper: BI-635 Introduction to Data Analysis using R Programming**

**CO1**  Students will know the importance and effectiveness of R language in analysis of biological data as compared to other platforms

**Activity 1**  Write any five R packages and their biological relevance

**Case Study**  Write any ten types of analyses of biological data that can be easily performed with R

| | Roll No | Test1 | Test2 | Total marks of attempted questions | Total Marks | Percentage | Score (3,2,1) | Target >= 60% (Y/N) |
|---|---|---|---|---|---|---|---|---|
| **Marks** | | 10 | 10 | 20 | 20 | | | |
| | 25201 | 9 | 9 | 18 | 20 | 90 | 3 | Y |
| | 25202 | 8 | 8 | 16 | 20 | 80 | 3 | Y |
| | 25203 | 8 | 9 | 17 | 20 | 85 | 3 | Y |
| | | | | | | Average | | 3 |

| %age | score |
|---|---|
| >=60 | 3 |
| 50-59.9 | 2 |
| 40-49.0 | 1 |

**CO2**  Students will be able to use various R objects to handle specific type of data

**T1**  Explain different R objects with examples

**A2**  Write a note of R objects to handle sequence data

| | Roll No | Test 1 | Assignment2 | Total marks of attempted questions | Total Marks | Percentage | Score (3,2,1) | Target |
|---|---|---|---|---|---|---|---|---|
| **Marks** | | 10 | 10 | 20 | 20 | | | |
| | 25201 | 9 | 9 | 18 | 20 | 90 | 3 | Y |
| | 25202 | 8 | 7 | 15 | 20 | 75 | 3 | Y |
| | 25203 | 9 | 9 | 18 | 20 | 90 | 3 | Y |
| | | | | | | Average | | 3 |

**CO3**  Students can be able to read big data from local as well as remote locations

**A1**  Download text data with different delimiters (comma, tab, semicolon etc.) and read them using R

**A2**  Download data from a remote server using download.file ()function in R

| | Roll No | Assignment1 | Assignment2 | Total marks of attempted questions | Total Marks | Percentage | Score (3,2,1) | Target |
|---|---|---|---|---|---|---|---|---|
| **Marks** | | 10 | 10 | | 20 | | | |
| | 25201 | 8 | 7 | 15 | 20 | 75 | 3 | Y |
| | 25202 | 7 | 8 | 15 | 20 | 75 | 3 | N |
| | 25203 | 8 | 7 | 15 | 20 | 75 | 3 | Y |

| | | | | | Average | | 3 |

**CO4** — Students can be able to statistically analyze biological data

**A1** — Download any eukaryotic genome and count different type of nucleotides and write results to a file

**A2** — Download gene expression data for cyclophilin gene family in wheat and generate a heatmap using R

| | Roll No | Test1 | Assignment1 | Total marks of attempted questions | Total Marks | Percentage | Score (3,2,1) | Target |
|---|---|---|---|---|---|---|---|---|
| Marks | | 10 | 10 | | 20 | | | |
| | 25201 | 9 | 9 | 18 | 20 | 90 | 3 | Y |
| | 25202 | 8 | 8 | 16 | 20 | 80 | 3 | N |
| | 25203 | 8 | 9 | 17 | 20 | 85 | 3 | Y |
| | | | | | | Average | | 3 |

**CO5** — Students will be able to write R scripts

**T1** — Write an R script to read RNAseq data, perform its Quality Control. Also draw the QC plots

**A1** — Using above pre-prcessed data generate Multidimensional scaling plots and heatmaps

| | Roll No | Test1 | Assignment1 | Total marks of attempted questions | Total Marks | Percentage | Score (3,2,1) | Target |
|---|---|---|---|---|---|---|---|---|
| Marks | | 10 | 20 | | 30 | | | |
| | 25201 | 9 | 18 | 27 | 30 | 90 | 3 | Y |
| | 25202 | 8 | 15 | 23 | 30 | 76.66666667 | 3 | N |
| | 25203 | 8 | 17 | 25 | 30 | 83.33333333 | 3 | Y |
| | | | | | | Average | | 3 |