U.S.N.										

B.M.S. College of Engineering, Bengaluru-560019

Autonomous Institute Affiliated to VTU

May / June 2025 Semester End Main Examinations

Programme: B.E. **Semester: VIII**
Branch: Biotechnology **Duration: 3 hrs.**
Course Code: 22BT8PEBDA **Max Marks: 100**
Course: Biological data analytics

Instructions: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

Important Note: Completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages. Revealing of identification, appeal to evaluator will be treated as malpractice.

			CO	*PO*	Marks
		UNIT - I			
1	a)	Compare various features of R with other programming languages R and justify why it is so popular for statistical analysis of big data.	*1*	*1*	**10**
	b)	Write R-syntax for control structures such as ***if else, loops, while loops and switch.***	*1*	*1*	**10**
		OR			
2	a)	Justify that R-is immensely useful for bioinformatics in the context of transcriptomics and proteomics by listing various R-packages and their utility	*1*	*1*	**10**
	b)	Write R-syntax for statistical analysis and data visualization.	*1*	*1*	**10**
		UNIT - II			
3	a)	R- Can be interfaced with other languages. Justify this statement with respect to C++.	*1*	*1,5*	**10**
	b)	List various time series concepts with biological applications and write R-syntax for each.	*1*	*1,5*	**10**
		OR			
4	a)	List various non-linear models with biological applications and write R-syntax for each.	*1*	*1,5*	**10**
	b)	What is the basic syntax of linear model? Implement R syntax for simple and multiple linear regression	*1*	*1,5*	**10**
		UNIT - III			
5	a)	Provide the list of various resources for genomics and write their utility	*1*	*1,5*	**10**

		b)	Implement any one of the R-packages to manage FASTAQ file, write commands for this procedure.	*1*	*1,5*	**10**
			OR			
	6	a)	Write R-Script for extracting Called Variants from VCF files using any one R-Package	*2*	*5*	**10**
		b)	Implement any R-package for processing/Reading WIG/BigWig files.	*2*	*5*	**10**
			UNIT - IV			
	7	a)	Design commands/step by step process for preparing a count matrix for RNA seq in R	*2*	*5*	**10**
		b)	Create steps for MA and Volcano plots using R commands	*2*	*5*	**10**
			OR			
	8	a)	Design series of steps for creating Heatmap of Top Differentially Expressed Genes and PCA plot using R	*3*	*1,5*	**10**
		b)	Develop steps for extracting the DESeq2 results using R	*3*	*1,5*	**10**
			UNIT - V			
	9	a)	Assume that you are working in proteomics research and would like to use various related databases. Can you List these databases and write briefly about their utility?	*3*	*1,5*	**10**
		b)	Justify the role of ML in proteomics with suitable examples.	*3*	*1,5*	**10**
			OR			
	10	a)	Write briefly on 1. Parent ion mass error 2. Filtering criteria	*3*	*1,5*	**10**
		b)	Assume that you are working with proteomics data and would like to use best available software packages for analyzing the data. What packages would you like to use? Justify the merits of five webtools.	*3*	*1,5*	**10**

****