

(Annexure F-4)

FACULTY OF COMPUTER SCIENCE & IT

SYLLABUS

of

Additional / Optional paper for specialization in Data Science

for

Bachelor of Computer Applications

(Semester III-IV)

(Under Continuous Evaluation System)

(12+3 System of Education)

Batch: 2022-25

Session: 2023-24



**The Heritage Institution
KANYA MAHA VIDYALAYA
JALANDHAR
(Autonomous)**

Kanya Maha Vidyalaya, Jalandhar (Autonomous)

SCHEME AND CURRICULUM OF EXAMINATIONS OF THREE YEAR DEGREE PROGRAMME

Bachelor of Computer Applications

Session 2023-24

Additional / Optional paper for Specialization in Data Science

Bachelor of Computer Applications Semester – III							
Course Code	Course Name	Course Type	Marks				Examination Time (in Hours)
			Total	Ext.		CA	
				L	P		
BCAL-3116	*Data Visualization	O	75	60	-	15	3
	Total		75				

Bachelor of Computer Applications Semester IV							
Course Code	Course Name	Course Type	Marks				Examination Time (in Hours)
			Total	Ext.		CA	
				L	P		
BCAM-4118	* Foundation of Statistical Computing	O	75	40	20	15	3+3
	Total		75				

Note:

O - Optional

*One additional/optional paper will be studied by the candidate if she opts for Specialization in Data Science

Bachelor of Computer Applications Semester- III
(Session 2023-24)
COURSE CODE: BCAL-3116
DATA VISUALIZATION

Course Outcomes:

After the completion of this course, the student will be able to:

CO1: Comprehend Importance and applications of data visualization.

CO2: Acquaint with categories of data visualization tools.

CO3: Apply basic and specialized visualization tools for representation of data.

CO4: Apply advanced data visualization tools.

Bachelor of Computer Applications Semester- III
(Session 2023-24)

COURSE CODE: BCAL-3116
DATA VISUALIZATION

Examination Time: 3 Hrs

Max. Marks: 75

Theory: 60

CA: 15

Instructions for Paper Setter -

Eight questions of equal marks (12 marks each) are to set, two in each of the four sections (A-D). Questions of Sections A-D should be set from Units I-IV of the syllabus respectively. Questions may be divided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each section. The fifth question may be attempted from any section

UNIT-I

Introduction: Introduction of Data Visualization, Meaning of Data Visualization, Importance of Data Visualization, Data Visualization applications, User psychology of Visualization, UX in Data Visualization.

UNIT-II

Gestalt principles of Data Visualization, Introduction to DIKW hierarchy, Goals of Data Visualization, Basic Visualization tools - Area Plots, Histograms, Bar Charts.

UNIT-III

Data Visualization tools: Introduction, characteristics, types, image and graphical visualization. Specialized Visualization tools: Pie Charts, Box Plots, Scatter Plots, Bubble Plots.

UNIT-IV

Advanced Visualization tools: Need, Application, Visualization of Maps, Storyboards in Visualization. Waffle Charts, Word Clouds, Seaborn and Regression Plots.

References / Textbooks:

1. E. Tufte, The Visual Display of Quantitative Information (2nd Edition), Graphics Press, 2001.
2. Herbert Jones, Data Science: The Ultimate Guide to Data Analytics, Data Mining, Data Warehousing, Data Visualization, Regression Analysis, Database Querying, Big Data for Business and Machine Learning for Beginners Kindle Edition, 2020.
3. E. Tufte, Envisioning Information, Graphics Press, 1990.
4. Andy Kirk, Data Visualisation: A Handbook for Data Driven Design SAGE Publications Ltd; First edition, 2016.

5. Kieran Healy, Data Visualization: A Practical Introduction Kindle Edition, Princeton University Press; First edition, 2018.

Note: The latest editions of the books should be followed.

Bachelor of Computer Applications Semester IV
(Session 2023-24)
COURSE CODE: BCAM-4118
FOUNDATION OF STATISTICAL COMPUTING

Course Outcomes:

After passing this course the student will be able to:

CO1: Comprehend basics of Statistical Computing and managing data structures like vector, matrix, etc.

CO2: Create, operate and manage lists and data frames.

CO3: Apply control and I/O statements for generating outputs.

CO4: Simulate various descriptive and analytical algorithms using R language along with their visualization.

Bachelor of Computer Applications Semester IV
(Session 2023-24)
COURSE CODE: BCAM-4118
FOUNDATION OF STATISTICAL COMPUTING

Examination Time: (3+3)Hrs

Max. Marks: 75
Theory: 40
Practical: 20
CA: 15

Instructions for Paper Setter -

Eight questions of equal marks (8 marks each) are to set, two in each of the four sections (A-D). Questions of Sections A-D should be set from Units I-IV of the syllabus respectively. Questions may be divided into parts (not exceeding four). Candidates are required to attempt five questions, selecting at least one question from each section. The fifth question may be attempted from any section

UNIT – I

Statistical Computing: Introduction, Role of Programming and Statistical Software. Data
Statistics: Sampling, Cumulative statistics, Statistics for Data frames, matrix objects and lists.

Introduction to R, Help functions in R, Vectors, Common Vector Operations, Using all and any function, subletting of vector. Creating matrices, Matrix operations, Applying Functions to Matrix Rows and Columns, Adding and deleting rows and columns.

UNIT - II

Lists, Creating lists, general list operations, Accessing list components and values, applying functions to lists, recursive lists

Creating Data Frames: Matrix-like operations in frames , Merging Data Frames, Applying functions to Data frames, Factors and Tables , factors and levels , Common functions used with factors , string operations

UNIT - III

Input/ Ouput:scan() , readline() Function, Printing to the Screen Reading and writing CSV and text file. Control statements: Loops, Looping Over Nonvector, Sets, if-else , writing user defined function, scope of the variable, R script file.

UNIT – IV

Graphics in R: Graph Syntax ((title, xlabel, ylabel, pch, lty, col.), Simple graphics (Bar, Multiple Bar, Histogram, Pie, Box-Plot, Scatter plot, qqplot), Low-level and High-Level plot functions. Using Analytical Algorithms (KNN, K-means, Naive Bayes) for Predictive analysis and Modelling.

References / Textbooks:

1. Andrie de Vries and JorisMeys, R Programming for Dummies, Wiley (2016), 2nd Edition.
2. SandipRakshit, R Programming for Beginners, McGraw Hill Education (2017), 1st Edition.
3. SandipRakshit, Statistics with R Programming, McGraw Hill Education (2018), 1st Edition.
4. Garrett Grolemond, Hands on Programming with R, O'Reilly (2014), 1st Edition
5. Mark Gardener, Beginning R: The Statistical Programming Language, Wiley (2013)
6. Tilman M. Davies, The Book of R: A first Course in Programming and Statistics, No Strach Press (2016), 1st Edition