(Annexure F-2)

# FACULTY OF COMPUTER SCIENCE & IT

**SYLLABUS**

**of**

**Additional / Optional paper for specialization in Data Science**

**for**
**Bachelor of Computer Applications**

**(Semester I-II)**

**(Under Credit Based Continuous Evaluation Grading System)**

**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

## Credit Based Continuous Evaluation Grading System (CBCEGS)

**Session 2023-24**

**Additional / Optional paper for Specialization in Data Science**

| **Bachelor of Computer Applications Semester – I** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Course Code** | **Course Name** | **Course Type** | **Hours per week** | **Credit** | | **Marks** | | | | **Examination Time (in Hours)** |
| | | | **L-T-P** | **L-T-P** | **Total** | **Total** | **Ext.** | | **CA** | |
| | | | | | | | **L** | **P** | | |
| **BCAL-1118** | ***Computational Data Science** | O | 3-1-0 | 3-1-0 | 4 | 100 | 80 | - | 20 | 3 |
| | **Total** | | | | 4 | 100 | | | | |

| **Bachelor of Computer Applications Semester – II** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Course Code** | **Course Name** | **Course Type** | **Hours per week** | **Credit** | | **Marks** | | | | **Examination Time (in Hours)** |
| | | | **L-T-P** | **L-T-P** | **Total** | **Total** | **Ext.** | | **CA** | |
| | | | | | | | **L** | **P** | | |
| **BCAL-2118** | ***Statistical Techniques for Data Science** | O | 3-1-0 | 3-1-0 | 4 | 100 | 80 | - | 20 | 3 |
| | **Total** | | | | 4 | 100 | | | | |

**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 I

**(Session 2023-24)**

**COURSE CODE: BCAL–1118**

# COMPUTATIONAL DATA SCIENCE

**Course Outcomes:**

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

CO1: Comprehend terminology associated with data and its processing.

CO2: Comprehend various types of functions in set theory.

CO3: Apply Algorithms of polynomial algebra to solve problems.

CO4: Apply various counting principles, permutations, combinations and averages to solve basic set of problems.

Bachelor of Computer Applications Semester I

**(Session 2023-24)**

**COURSE CODE: BCAL–1118**

# COMPUTATIONAL DATA SCIENCE

**L-T-P: 3-1-0** **Max. Marks: 100**
**Credit: 4** **Theory: 80**
**Examination Time: 3 Hrs.** **CA: 20**

**Instructions for Paper Setter -**

Eight questions of equal marks (16 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

**Data Processing:** Basic Terminology of Data, Types of Data, Information and Knowledge, Preprocessing the Data, Data cleaning, Data transformation, Data reduction.

Introduction to Data Science, Evolution of Data science, Need of Data Science, Components of Data Science, Application Areas.

## UNIT II

**Functions**: Functions and their types, Quadratic Functions and Equations, Inverse Function, Logarithmic Functions and Equations.

## UNIT III

**Algebra of Polynomials**: Addition, Subtraction, Multiplication and Division Algorithms
**Graphs of Polynomials**: X-intercepts, multiplicities, end behavior and turning points, Graphing & Polynomial Creation.

## UNIT IV

**Basic Principles of Counting and Factorial Concepts**: Addition rule of counting, Multiplication rule of counting, Factorials.
Permutation and Combination.
**Measures of Central Tendency:** Mean, Median and Mode

**References/Textbooks:**

1. Patricia Pulliam Phillips, Cathy A. Stawarski, “Data Collection: Planning for and Collecting All Types of Data”, Wiley Publisher, First Edition, 2008.

2. Roger Sapsford, Victor Jupp, “Data Collection-and Analysis”, Second Edition, Sage Publishing, 2006.
3. Kenneth Rosen, “Discrete Mathematics and Its Applications”, Tata McGraw Hill, 7th Edition
4. Anshuman Sharma, Fundamentals of Numerical Methods and Statistical techniques, Lakhanpal Publications (2016)

Bachelor of Computer Applications Semester II

**(Session 2023-24)**

**COURSE CODE: BCAL–2118**

**STATISTICAL TECHNIQUES FOR DATA SCIENCE**

**Course Outcomes:**

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

CO1: Comprehend the key terminology of descriptive statistics and frequency distribution

CO2: Comprehend the basic Probability terms and their usage.

CO3: Formulate hypothesis for basic problems and perform testing.

CO4: Implement statistical techniques like Chi Square test and Analysis of variance.

Bachelor of Computer Applications Semester II

**(Session 2023-24)**

**COURSE CODE: BCAL–2118**

**STATISTICAL TECHNIQUES FOR DATA SCIENCE**

**L-T-P: 3-1-0** **Max. Marks: 100**
**Credit: 4** **Theory: 80**
**Examination Time: 3 Hrs** **CA: 20**

**Instructions for Paper Setter -**

Eight questions of equal marks (16 mark 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 to Statistics:** Basic terminology, variables: discrete and continuous.

**Introduction to descriptive Statistics:** Types of data, levels of measurement, categorical variables and numerical variables. Introduction to Frequency distribution.

**Measures of Dispersion:** Quartile Deviation, Mean Deviation, Standard Deviation and Coefficient of Variance.

**UNIT – II**

**Introduction to Asymmetry:** Moments, Kurtosis and Skewness.

**Probability:** Meaning, Basic concepts, Events, Properties of Probability, Conditional Probability.

**UNIT - III**

**Probability:** Addition Theorem, Multiplication Theorem and Bayes' Theorem.

**Introduction to Inferential statistics:** Concept of a sample and a population, need of sampling.

**Hypothesis Testing:** Null and Alternate Hypothesis, Type 1 and Type 2 errors, Confidence intervals. Chi square test.

**UNIT - IV**

**ANOVA** - one way and two way.

**Data Analysis Tools in Spreadsheets:** Regression Analysis, Correlation Analysis, Covariance Analysis, ANOVA Analysis.

**References/Textbooks:**


1. S.P Gupta, Statistical Methods, Sultan Chand & Sons (2012)
2. B. L. Agarwal, Statistics For Professional Courses, CBS Professional (2011)

3. Anshuman Sharma, Fundamentals of Numerical Methods and Statistical techniques, Lakhanpal Publications (2016)
4. Stephen L. Nelson, Excel Data Analysis for Dummies, Wiley Publications (2013)