PARVATHANENI BRAHMAYYA
SIDDHARTHA COLLEGE OF ARTS & SCIENCE
Autonomous
Siddhartha Nagar, Vijayawada–520010
Re-accredited at 'A+' by the NAAC

23STMIL122: Descriptive Statistics for Data Analytics

Offered to: All UG Programs **Course Type:** Minor 1 (Core -TH)

Year of Introduction: 2023-24 **Year of offering: 2023 - 2024**

Semester: II 60 Hrs **Credits: 3**

Course Prerequisites: Basic knowledge in Mathematics.

Course Description:
This course helps the students to familiarize with the ways in which we talk about descriptive statistics, uncertainty and estimate their situations in which probability arises. Also this course aims at providing basic knowledge about theoretical and applications of attributes.

Course Objectives:

1) To compute various measures of central tendency, dispersion, skewness and kurtosis.
2) To find the probabilities of events.
3) To get the knowledge regarding qualitative factors

Course Outcomes:		
Course Outcome	Upon successful completion of this course, students should have the knowledge and skills to:	**PO Mapping**
CO 1	To impart knowledge on Statistical concepts like Data Collection and Classification.	PO1
CO 2	To learn different types measures of central tendency and dispersion.	PO2
CO3	Students will be able to draw the descriptive statistics for the data and interpret the data with the appropriate diagrams and graphs.	PO3
CO 4	Understand the basic concepts of probability and to find probabilities of various events.	PO3
CO 5	Get the knowledge in respect of usage in day-to-day life in decision making in the face of uncertainty	PO3

CO-PO MATRIX							
COURSE CODE	**CO-PO**	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**
23STMIL122	**CO1**	**3**					
	CO2		**3**				
	CO3			**3**			
	CO4			**3**			
	CO5			**3**			

Syllabus

Unit	Learning Units	Lecture Hours
I	Measurement Scales – Nominal, Ordinal, Ratio and Interval. Frequency distribution and types of frequency distributions, forming a frequency distribution. Diagrammatic representation of data – Histogram, Simple Bar, Multiple bar and Pie with simple problems. Graphical representation of data: Histogram, frequency polygon and Ogives with simple problems.	12
II	**Measures of Central Tendency (MCT)** Objectives of averages, Characteristics of a good average. Arithmetic mean, Geometric mean, Harmonic mean, Median and Mode – merits, demerits, properties and applications. **Measures of Dispersion** Significance of measures of dispersion, Characteristic of an ideal measure of dispersion, Absolute and relative measures of dispersion – range, quartile deviation, mean deviation, variance and standard deviation – merits, demerits, properties and applications	12
III	**Moments:** Central and non-central moments and their inter-relationships, Sheppard's corrections for moments for grouped data and problems. **Skewness:** Definition, measures of skewness by Karl Pearson's, Bowley's formulae and based on moments and problems. **Kurtosis:** Definition, measures of kurtosis based on moments and problems.	12
IV	**Probability – I** Terminology - Random experiments, trial, sample space, mutually exclusive , exhaustive , equally likely, favorable, conditional and independent events. Definitions-Mathematical, Statistical and Axiomatic definitions of probabilities. Addition law of probabilities for two, three and n events. Boole's inequalities and problems.	12
V	**Probability – II** Conditional Probability- multiplication law of probability for two and n events. Pairwise independent events and conditions for mutual independence of n events and Baye's theorem and its applications. Introductory concept of Geometric probability.	12

Text Book:

Fundamentals of Mathematical Statistics, 11th Edition, 2010, S. C. Gupta and V. K. Kapoor, Sultan Chand & Sons, New Delhi

Reference Books:

1. B.A/B.Sc. Second Year Statistics(2010) , Telugu Akademi, Hyderabad.
2. Mathematical Statistics with Applications, 2009, K.M.Ramachandran and Chris P.Tsokos Academic Press(Elsevier), Haryana .
3. Probability and Statistics, Volume I & II, D. Biswas, New central book Agency (P) Ltd, NewDelhi.
4. An outline of Statistical theory, Volume II,3rd Edition,2010(with corrections) A.M.Goon,M.K. Gupta, B.Dasgupta ,The World Press Pvt.Ltd., Kolakota.

Sanjay Arora and BansiLal:. New Mathematical Statistics, SatyaPrakashan , New Delhi.

PARVATHANENI BRAHMAYYA
SIDDHARTHA COLLEGE OF ARTS & SCIENCE
Autonomous
Siddhartha Nagar, Vijayawada–520010
Re-accredited at 'A+' by the NAAC

23STMIL122: Descriptive Statistics for Data Analytics

Minor 1 Semester II

Time: 3 hours Maximum Marks: 70

Section – A

Answer the following **5 x 4M = 20Marks**

1. (a) What are different sources of Primary data? (L-2, CO-1)

   (Or)

   (b) Name two kinds of statistical data and describe them in brief. (L-2, CO-1)

2. (a) Calculate Arithmetic mean for the following data (L-2, CO-2)

Class intervals	0-8	8-16	16-24	24-32	32-40	40-48
Frequency	8	7	16	24	15	7

   (OR)

   (b) Characteristics of ideal measures of central tendency. (L-2, CO-2)

3. (a) Show that for discrete distributions$\beta_1 > 1$. (L-2, CO-3)

   (OR)

   (b) The first four moments of a distribution about the value 5 are -4, 22,-117 and 560. Find the corresponding moments about the mean and also comment on the nature of the data. (L-2, CO-3)

4. (a) State and prove addition theorem of probability for two events. (L-1, CO-4)

   (OR)

   (b) Define axiomatic definition of probability. (L-1, CO-4)

5. (a) If A and B are independent events, then prove that (L-1, CO-5)
   (i) $\bar{A}$ and B (ii) $\bar{A}$ and $\bar{B}$ are also independent.

   (OR)

   (b) State and prove multiplication theorem of probability. (L-1, CO-5)

Section – B

Answer the following **5 x 10M = 50Marks**

6. a) Draw the histogram for table depicts the number of students of a class engaged in any one of the sports. Note that the number of students is actually the frequency. (L-2, CO-1)

Sports	Archery	Cycling	Power lifting	Swimming	Snooker	Table Tennis	Skate boarding
Frequency	8	12	13	15	14	10	9

   (OR)

   b) Draw Ogive curve to the following data and also obtain median through Ogives

Wages (in Rs.)	50-60	60-70	70-80	80-90	90-100	100-110	110-120	120-130
No. Of workers	15	20	34	50	70	45	26	10

(L-2, CO-1)

7. (a) Calculate Mean, Median and Mode for the following data (L-3, CO-2)

Class intervals	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Frequency	15	20	34	40	50	30	10

(OR)

(b) Calculate Standard deviation and coefficient of variation for the following table giving the age Distribution of 542 members of a club:

Age (in Years)	20-30	30-40	40-50	50-60	60-70	70-80	80-90
No. Of Members	3	61	132	153	140	51	2

(L-3, CO-2)

8. (a) Define moments. Establish the relationship between the moments about mean (Central moments) in terms of moments about any arbitrary point and vice versa. (L-3, CO-3)

(OR)

(b) The scores in statistics of 250 candidates appearing at an examination have mean = 39.72, variance = 97.80, 3rd central moment and 4th central moments are -114.18 and 28,396.14. It was later found on scrutiny that the score 61 of a candidate has been wrongly recorded as 51. Make necessary corrections in the given values of the mean and central moments. (L-3, CO-3)

9. (a) State and prove Boole's inequality. (L-3, CO-4)

(OR)

(b) State and prove the addition theorem of probability for n events. (L-3, CO-4)

10. (a) For two events A and B, prove that (L-3, CO-5)

(i) $P(\bar{A} \cap B) = P(B) - P(A \cap B)$ (ii) $P(A \cap \bar{B}) = P(A) - P(A \cap B)$

(iii) If $B \subset A$ then $P(A \cap \bar{B}) = P(A) - P(B)$ (iv) If $A \subset B$ then $P(\bar{A} \cap B) = P(B) - P(A)$

(OR)

(b) State and prove Baye's theorem of probability. (L-3, CO-5)

***