

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

BSc DEGREE EXAMINATION MAY 2026
(Second Semester)

Branch – COMPUTER SCIENCE DATA ANALYTICS

STATISTICAL DATA ANALYSIS

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer ALL questions

ALL questions carry EQUAL marks (10 × 1 = 10)

Module No.	Question No.	Question	K Level	CO
1	1	In a bar diagram, the base line is (i) Horizontal (ii) vertical (iii) False base line (iv) Diagonal	K1	CO1
	2	The mode of the distribution of value 5, 7,9,9,8,5,6,8,7,7,5,7,9,2,7, is (i) 5 (ii) 8 (iii) 9 (iv)7	K2	CO1
2	3	The range of simple correlation coefficient is (i) 0 to ∞ (ii) - ∞ to ∞ (iii) 0 to 1 (iv) -1to1	K1	CO2
	4	The regression coefficient β_{yx} is theof the regression line (i) Slope (ii) Intersect (iii) Horizontal (iv) parallel	K2	CO2
3	5	The probability of an event cannot be (i) Equal to zero (ii) greater than zero (iii) Equal to one (iv) Less than zero	K1	CO3
	6	The parameters of the binomial distribution are: (i) n and p (ii) p and q (iii) np and nq (iv) np and npq	K2	CO3
4	7	The ranges of 't' distribution is (i) 0 to ∞ (ii) - ∞ to ∞ (iii) 0 to 1 (iv) -1to1	K1	CO4
	8	When standard deviation of the population is known (i) $S.E_{\bar{X}} = \frac{\sigma_p}{\sqrt{n}}$ (ii) $S.E_{\bar{X}} = \frac{\sigma_p}{\sqrt{2n}}$ (iii) $S.E_{\bar{X}} = \frac{\sigma_p^2}{\sqrt{n}}$ (iv) $S.E_{\bar{X}} = \frac{\sigma_p}{\sqrt{n}}$	K2	CO4
5	9	If all frequencies of classes are same, the value of χ^2 is (i) 1 (ii) 0 (iii) ∞ (iv) 2	K1	CO5
	10	Which of the following is the assumption of 'F' test (i) Equality (ii) Nullity (iii) Normality (iv) Error	K2	CO5

Cont...

SECTION - B (35 Marks)Answer **ALL** questions**ALL** questions carry **EQUAL** Marks

(5 × 7 = 35)

Module No.	Question No.	Question	K Level	CO																											
1	11.a.	Represent the following data by a suitable diagram <table><tr><td>Year</td><td>2013</td><td>2014</td><td>2015</td><td>2016</td><td>2017</td><td>2018</td><td>2019</td><td>2020</td></tr><tr><td>Imports</td><td>400</td><td>450</td><td>500</td><td>620</td><td>560</td><td>460</td><td>500</td><td>540</td></tr><tr><td>Exports</td><td>220</td><td>235</td><td>385</td><td>420</td><td>420</td><td>380</td><td>360</td><td>400</td></tr></table>	Year	2013	2014	2015	2016	2017	2018	2019	2020	Imports	400	450	500	620	560	460	500	540	Exports	220	235	385	420	420	380	360	400	K4	CO1
	Year	2013	2014	2015	2016	2017	2018	2019	2020																						
	Imports	400	450	500	620	560	460	500	540																						
Exports	220	235	385	420	420	380	360	400																							
	(OR)																														
11.b.	Calculate the standard deviation of the following data <table><tr><td>X</td><td>6</td><td>9</td><td>12</td><td>15</td><td>18</td></tr><tr><td>f</td><td>7</td><td>12</td><td>13</td><td>10</td><td>8</td></tr></table>	X	6	9	12	15	18	f	7	12	13	10	8																		
X	6	9	12	15	18																										
f	7	12	13	10	8																										
2	12.a.	List the types of correlation.	K4	CO2																											
		(OR)																													
	12.b.	For the data given below calculate the rank correlation coefficient <table><tr><td>X</td><td>21</td><td>36</td><td>49</td><td>37</td><td>25</td></tr><tr><td>Y</td><td>47</td><td>40</td><td>37</td><td>42</td><td>43</td></tr></table>			X	21	36	49	37	25	Y	47	40	37	42	43															
X	21	36	49	37	25																										
Y	47	40	37	42	43																										
3	13.a.	A bag contains 30 balls numbered from 1 to 30. One ball is drawn at random. Find the probability that the number of the ball drawn will be a multiple of (a) 5 or 7 and (b) 3 or 7.	K3	CO3																											
		(OR)																													
	13.b.	A coin is tossed 6 times. What is the probability of obtaining four or more heads?																													
4	14.a.	In a simple random sample of 600 men taken from a big city, 400 found to be smokers. In another simple random sample of 900 men taken from another city 450 are smokers. Do the data indicate that there is a significant difference in the habit of smoking in the two cities?	K5	CO4																											
		(OR)																													
	14.b.	The life time of electric bulbs for a random sample of 10 from a large consignments gave the following data <table><tr><td>Item</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr><tr><td>Life in '000 hours</td><td>4.2</td><td>4.6</td><td>3.9</td><td>4.1</td><td>5.2</td><td>3.8</td><td>3.9</td><td>4.3</td><td>4.4</td><td>5.6</td></tr></table> Can we accept the hypothesis that the average life time of bulbs is 4,000 hours.			Item	1	2	3	4	5	6	7	8	9	10	Life in '000 hours	4.2	4.6	3.9	4.1	5.2	3.8	3.9	4.3	4.4	5.6					
Item	1	2	3	4	5	6	7	8	9	10																					
Life in '000 hours	4.2	4.6	3.9	4.1	5.2	3.8	3.9	4.3	4.4	5.6																					
5	15.a.	In a test given to two groups of students drawn from two normal populations, the marks obtained were obtained as follows. Group A : 18 20 36 50 49 36 34 49 41 Group B : 29 28 26 35 30 44 46 Examine at 5% level, whether the two population have the same variance.	K5	CO5																											
		(OR)																													
	15.b.	On the basis of information noted below, find out whether the new treatment is comparatively superior to the conventional one. <table><tr><td></td><td>Favorable</td><td>Non favorable</td><td>Total</td></tr><tr><td>Conventional</td><td>40</td><td>70</td><td>110</td></tr><tr><td>New</td><td>60</td><td>30</td><td>90</td></tr><tr><td>Total</td><td>100</td><td>100</td><td>N=200</td></tr></table>				Favorable	Non favorable	Total	Conventional	40	70	110	New	60	30	90	Total	100	100	N=200											
	Favorable	Non favorable	Total																												
Conventional	40	70	110																												
New	60	30	90																												
Total	100	100	N=200																												

Cont...

SECTION -C (30 Marks)
Answer ANY THREE questions
ALL questions carry EQUAL Marks (3 × 10 = 30)

Module No.	Question No.	Question	K Level	CO																								
1	16	Calculate the Arithmetic mean, Median, Mode from the following data <table><tr><td>Marks</td><td>11 – 20</td><td>21 – 30</td><td>31 – 40</td><td>41 – 50</td><td>51 – 60</td><td>61 – 70</td><td>71 – 80</td></tr><tr><td>No of Students</td><td>42</td><td>38</td><td>120</td><td>84</td><td>48</td><td>36</td><td>31</td></tr></table>	Marks	11 – 20	21 – 30	31 – 40	41 – 50	51 – 60	61 – 70	71 – 80	No of Students	42	38	120	84	48	36	31	K4	CO1								
Marks	11 – 20	21 – 30	31 – 40	41 – 50	51 – 60	61 – 70	71 – 80																					
No of Students	42	38	120	84	48	36	31																					
2	17	Calculate the correlation coefficient from the following data <table><tr><td>X</td><td>10</td><td>12</td><td>18</td><td>8</td><td>13</td><td>20</td><td>22</td><td>15</td><td>5</td><td>17</td></tr><tr><td>Y</td><td>88</td><td>90</td><td>94</td><td>86</td><td>87</td><td>92</td><td>96</td><td>94</td><td>88</td><td>85</td></tr></table>	X	10	12	18	8	13	20	22	15	5	17	Y	88	90	94	86	87	92	96	94	88	85	K4	CO2		
X	10	12	18	8	13	20	22	15	5	17																		
Y	88	90	94	86	87	92	96	94	88	85																		
3	18	The following data show the number of seeds germinating out of 10 on damp filter for 80 sets of seeds. Fit a binomial distribution to this data. <table><tr><td>x</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr><tr><td>f</td><td>6</td><td>20</td><td>28</td><td>12</td><td>8</td><td>6</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr></table>	x	0	1	2	3	4	5	6	7	8	9	10	f	6	20	28	12	8	6	0	0	0	0	0	K3	CO3
x	0	1	2	3	4	5	6	7	8	9	10																	
f	6	20	28	12	8	6	0	0	0	0	0																	
4	19	Two types of drugs were used on 5 and 7 patients for reducing their weight. Drug A was imported and drug B indigenous. The decrease in the weight after using the drugs for six months was follows: <table><tr><td>Drug A</td><td>10</td><td>12</td><td>13</td><td>11</td><td>14</td><td></td><td></td></tr><tr><td>Drug B</td><td>8</td><td>9</td><td>12</td><td>14</td><td>15</td><td>10</td><td>9</td></tr></table> Is there a significant difference in efficacy of the two drugs? If not, which drug should you buy?	Drug A	10	12	13	11	14			Drug B	8	9	12	14	15	10	9	K5	CO4								
Drug A	10	12	13	11	14																							
Drug B	8	9	12	14	15	10	9																					
5	20	Three samples below have been obtained from normal population with equal variances .test the hypothesis that the samples means are equal. <table><tr><td colspan="3">Samples.</td></tr><tr><td>8</td><td>7</td><td>12</td></tr><tr><td>10</td><td>5</td><td>19</td></tr><tr><td>7</td><td>10</td><td>13</td></tr><tr><td>14</td><td>9</td><td>12</td></tr><tr><td>11</td><td>9</td><td>14</td></tr></table> The value of F at 5% level of significance is 3.88	Samples.			8	7	12	10	5	19	7	10	13	14	9	12	11	9	14	K5	CO5						
Samples.																												
8	7	12																										
10	5	19																										
7	10	13																										
14	9	12																										
11	9	14																										

Z-Z-Z

END

