

--	--	--	--	--	--	--	--	--	--

Seventh Semester B.E. Degree Examination, Feb./Mar. 2022

Introduction to Big Data Analytics

Time: 3 hrs.

Max. Marks: 100

Note: Answer any FIVE full questions, choosing ONE full question from each module.

Module-1

- 1 a. Define model. Explain various steps of modeling process. (08 Marks)
- b. Explain different types of data with an example. (10 Marks)
- c. Explain graphical model. (02 Marks)

OR

- 2 a. Explain Box plot and Histograms with a neat diagram. (10 Marks)
- b. Explain stacked format and unstacked formats. (05 Marks)
- c. Explain correlation and covariance. (05 Marks)

Module-2

- 3 a. Explain the following with an example. (10 Marks)
 - i) Mutually exclusive events
 - ii) Conditional probability
 - iii) Equally likely events
 - iv) Subjective probability
 - v) Objective probability.
- b. Describe the summary measures of probability distribution. (08 Marks)
- c. Assume 10, 20, 30 and 40 are possible values of random variable X, with probabilities 0.15, 0.25, 0.35 and 0.25. Find $P(X \leq 30)$. (02 Marks)

OR

- 4 a. What is a density function? Explain Normal Distribution. (10 Marks)
- b. Explain the Microsoft excel functions for below probability distributions : (10 Marks)
 - i) Normal Distribution
 - ii) Binomial Distribution
 - iii) Poisson Distribution
 - iv) Exponential Distribution

Module-3

- 5 a. Construct the decision tree for the given data in Table Q5(a) :

		Outcome		
		O ₁	O ₂	O ₃
Decision	D1	10	10	10
	D2	-10	20	30
	D3	-30	30	80

Table Q5(a)

- b. Explain the various conventions used in Decision Tree. (10 Marks)
- c. Explain Baye's Rules. (05 Marks)
- d. Explain: (05 Marks)
 - i) Utility function
 - ii) Exponential utility.

OR

- 6 a. Explain different methods for selecting random samples. (10 Marks)
b. Explain the sources of estimation Errors. (10 Marks)

Module-4

- 7 a. Explain t distribution with respect to sampling. (10 Marks)
b. Give the applications of comparisons of means in Business. (05 Marks)
c. Explain sample size selection. (05 Marks)

OR

- 8 a. Explain how to find the significance of sample evidence from P-values. (07 Marks)
b. Explain Chi-square goodness of fit test for Normality. (10 Marks)
c. Explain different types of Errors. (03 Marks)

Module-5

- 9 a. Explain simple Linear regression using least square estimation. (08 Marks)
b. Give the characteristics of Multiple Regression. (05 Marks)
c. Explain Non-linear transformation for examining the variables. (07 Marks)

OR

- 10 a. How to analyze different sources of variation using ANOVA table. (08 Marks)
b. Give the guidelines for Including/Excluding variable in a Regression equation. (05 Marks)
c. Explain how to predict the value of the dependent variable for new observations. (07 Marks)

* * * * *