

Roll No.

Total No. of Pages : 02

Total No. of Questions : 17

**MBA (2018 Batch) (Sem.-3)**  
**DATA SCIENCES USING R**  
Subject Code : MBA-962-18  
M.Code : 77058

Time : 3 Hrs.

Max. Marks : 60

**INSTRUCTIONS TO CANDIDATES :**

1. SECTION-A contains EIGHT questions carrying TWO marks each and students has to attempt ALL questions.
2. SECTION-B consists of FOUR Subsections : Units-I, II, III & IV. Each Subsection contains TWO questions each carrying EIGHT marks each and student has to attempt any ONE question from each Subsection.
3. SECTION-C is COMPULSORY and consist of ONE question carrying TWELVE marks.

**SECTION-A**

**Explain briefly :**

Q1. Big Data

Q2. Mean, Median, Mode

Q3. Regression

Q4. SVM

Q5. Deep Learning

Q6. Decision Tree

Q7. Validation

Q8. Data Mining

## SECTION-B

### UNIT-I

- Q9. Differentiate between supervised and unsupervised learning.
- Q10. Explain standard deviation and variance using R software.

### UNIT-II

- Q11. What is decision tree? Where they are used? Explain with example.
- Q12. Explain the concept of linear, multiple and logistic regression.

### UNIT-III

- Q13. Explain the process of clustering with k-medoids algorithm.
- Q14. Write a note on hierarchical and density-based spatial clustering.

### UNIT-IV

- Q15. How evaluation and validation is achieved using cross-validation?
- Q16. Write a note on big data related techniques.

## SECTION-C

- Q17. A librarian purchases 500 books for the library. How data mining techniques (clustering and classification) will be used to arrange the books into racks?

**NOTE : Disclosure of Identity by writing Mobile No. or Making of passing request on any page of Answer Sheet will lead to UMC against the Student.**