

Total No. of Questions : 5]

SEAT No. :

PA-1971

[Total No. of Pages : 2

[5954]-306

S.Y. B.B.A. (C.A.)

CA- 305 : BIG DATA

(2019 CBCS Pattern) (Semester-III)

*Time : 2½ Hours]*

*[Max. Marks : 70*

*Instruction to the candidates:*

- 1) *All questions are compulsory.*
- 2) *Figures to right indicate marks.*

**Q1)** Attempt any EIGHT of the following.

**[16]**

- a) What is big data?
- b) What is data manipulation?
- c) What is data science?
- d) What is statistical Inference?
- e) Enlist the stages of data science?
- f) Define Machine Learning.
- g) Define SVM?
- h) What is the use of histogram?
- i) What is data analysis?
- j) What is the use of themes?

**Q2)** Attempt any FOUR of the following.

**[16]**

- a) Explain different types of data analytics.
- b) Give advantages and Disadvantages of Machine Learning.
- c) Explain the process of data analysis.
- d) Explain probability distribution modeling.
- e) Explain applications of big data.

**P.T.O.**

**Q3)** Attempt any FOUR of the following. **[16]**

- a) State advantages and disadvantages of SVM.
- b) Explain Data frame with example.
- c) Explain types of regression models.
- d) What is histogram? Explain with example in R.
- e) Explain functions included in “dplyr” package

**Q4)** Attempt any FOUR of the following. **[16]**

- a) Explain Naive Bayes with the help of example.
- b) What is data visualization? Explain with example in R.
- c) Write a R program to accept temperatures in Fahrenheit (F) and print it in Celsius (C).
- d) Accept three dimensions length (l), breadth (b) and height (h) of a cuboid and print its volume.
- e) Write a R program accept any year as input and check whether the year is a leap year or not.

**Q5)** Write a short note on Any TWO of the following. **[6]**

- a) Tools used in Big Data.
- b) Advantages of Big data.
- c) Advantages and Disadvantages of EM algorithms.

