Hall Ticket No ⬚⬚⬚⬚⬚⬚⬚⬚⬚⬚ Course Code: ACDCO1

INSTITUTE OF AERONAUTICAL ENGINEERING
(Autonomous)
Dundigal, Hyderabad - 500 043

B.Tech III SEMESTER END EXAMINATIONS (REGULAR / SUPPLEMENTARY) - FEBRUARY 2023
Regulation: UG20
PRINCIPLES OF DATA SCIENCE

Time: 3 Hours   **CSE (DATA SCIENCE)**   Max Marks: **70**

Answer ALL questions in Module I and II
Answer ONE out of two questions in Modules III, IV and V
All Questions Carry Equal Marks
All parts of the question must be answered in one place only

MODULE – I

1. (a) Outline the four levels of data with an example. Discuss the five essential steps to perform data science. [BL: Understand| CO: 1|Marks: 7]

   (b) Categorize the following variables and Justify whether they are

   i) Quantitative or qualitative data

   ii) Discrete or continuous data

   - Height of student
   - Eye color
   - Room temperature
   - Weight of newborn babies
   - Ducks in a pond
   - Rolling a dice
   - Bank account balance [BL: Apply| CO: 1|Marks: 7]

MODULE – II

2. (a) Interpret the meaning of each basic data types in R programming by providing a suitable variable. [BL: Understand| CO: 2|Marks: 7]

   (b) Consider the data given in Table 1

Table 1

S.No	Roll	Name	Mark
1	101	Abi	87.5
2	105	Banu	78.0
3	108	Sibi	98.5

   i) Create the dataframe

   ii) Access the name column using index value and column name

   iii) Add new rows to the existing Data Frame

      iv) Add new columns in a data frame
      v) Determine the size of a data frame
      vi) Find the number of columns and rows

[BL: Apply| CO: 2|Marks: 7]

MODULE – III

3. (a) Infer the knowledge about the operations, notation and Venn diagrams of set theory.

[BL: Understand| CO: 3|Marks: 7]

   (b) Two vectors are given by $\vec{a} = 2i + j + k$ and $\vec{b} = i + j + k$. Elaborate the following operation and determine it.

i) $\vec{a}\,\vec{b}$ ii) $\vec{a}\,\vec{b}$ iii) $5\vec{a} + 2\vec{b}$ iv) $\vec{a} + \vec{b}$ v) $\vec{a} - \vec{b}$ [BL: Apply| CO: 3|Marks: 7]

4. (a) Discuss about the different types of random variables. State the properties of probability density function

[BL: Understand| CO: 4|Marks: 7]

   (b) State Bayes theorem. In a neighbourhood, 90% children were falling sick due flu and 10% due to measles and no other disease. The probability of observing rashes for measles is 0.95 and for flu is 0.08. If a child develops rashes, find the child's probability of having flu.

[BL: Apply| CO: 4|Marks: 7]

MODULE – IV

5. (a) Explain the different methods for drawing samples from data in statistics for data science.

[BL: Understand| CO: 5|Marks: 7]

   (b) Data collection is an important aspect of research. Let's consider an example of a mobile manufacturer, company X, which is launching a new product variant. To conduct research about features, price range, target market, competitor analysis etc. data has to be collected from appropriate sources. The marketing team can conduct data collection activities by conducting survey and make questionnaire Design about it.   [BL: Apply| CO: 5|Marks: 7]

6. (a) Write about confidence interval. Summarize the purpose of using it and also how to calculate it.

[BL: Understand| CO: 5|Marks: 7]

   (b) A principal at a certain school claims that the students in his school are above average intelligence. A random sample of thirty students IQ scores have a mean score of 112.5. Make use of one Sample Z hypothesis testing to answer the question: "Is there sufficient evidence to support the principal's claim?". The mean population IQ is 100 with a standard deviation of 15.

[BL: Apply| CO: 5|Marks: 7]

MODULE – V

7. (a) Illustrate the process of visualizing the data using scatter plots and barchart.

[BL: Understand| CO: 6|Marks: 7]

   (b) Table 2 gives the lifetime of 400 neon lamps. Draw the histogram.

Table 2

Lifetime (in hours)	Number of lamps
300 – 400	14
400 – 500	56
500 – 600	60
600 – 700	86
700 – 800	74
800 – 900	62
900 – 1000	48

[BL: Apply| CO: 6|Marks: 7]

8. (a) Compare and contrast correlation and causation. Summarize the various effective data presentation techniques. [BL: Understand| CO: 6|Marks: 7]

(b) How do you evaluate visualization? Describe Simpson's paradox and how does it pertain to confounding? [BL: Apply| CO: 6|Marks: 7]

− ∘ ∘ ◯ ∘ ∘ −