### Paper / Subject Code: 37471 / Data Analytics and Visualization

May 15, 2024 02:30 pm - 05:30 pm 1T01876 - T.E. Computer Science and Engineering (Artificial Intelligence and Machine Learning) (Choice Based) (R-19- 'C' Scheme)SEMESTER - VI / 37471 - Data Analytics and VisualizationQP CODE: 10055196

# (3 Hours)

(Total Marks: 80)

N.B.: 1. Question No. 1 is compulsory.

- 2. Answer any three out of the remaining questions.
- 3. Assume suitable data if necessary.
- 4. Figures to the right indicate full marks.

# Q1. Attempt any FOUR

**f20**1

- [A] List and explain different key roles for successful data analytics?
- [B] What is Stepwise regression? Explain its types.
- [C] Explain Term Frequency-Inverse Document Frequency (TF-IDF) with a suitable example.
- [D] Difference between Matplotlib and Seaborn library.
- [E] Explain components of time series?

#### Q2. Attempt the following

[20]

- [A] Explain different phases in data analytics lifecycle.
- [B] Explain ARIMA model in detail. Also state its Pros and Cons.

### Q3. Attempt the following

[20]

- [A] Explain in detail seven practice areas of text analytics.
- [B] Explain different types of data visualisation in Python programming language.

#### Q4. Attempt the following

[20]

[A] The number of bacterial cells (y) per unit volume in a culture at different hours (x) is given below:

	X	0		2	3	48	5	6	7	8	9
3	y	43	46	82	98	123	167	199	213	245	272

Fit lines of regression of y on x and x on y. Also, estimate the number of bacterial cells after 15 hours.

55196