

Total No. of Questions – [06]

Total No. of Printed Pages:[01]

G.R. No.	
----------	--

PAPER CODE	P122-231 PSE/ESE
------------	------------------

May 2022 / INSEM+ENDSEM
F. Y. M. TECH. (COMPUTER ENGINEERING)
(SEMESTER – II)
COURSE NAME: DATA SCIENCES
COURSE CODE: CSPA12201
(PATTERN 2020)

Time: [3 Hours]

[Max. Marks: 60]

(*) Instructions to candidates:

- 1) All Questions are compulsory
- 2) Figures to the right indicate full marks.
- 3) Use of scientific calculator is allowed
- 4) Use suitable data where ever required

- Q.1 Justify the need of exploratory data analysis important in data science. [10]
- Q.2 Identify the suitable algorithm to implement plagiarism detection software. [10]
- Q.3 Identify and explain the method to finding and recommending suitable items that would be liked and selected by users. [10]
- Q4. In an application you have given text file and which method you will use to find out important word? [10]
- Q5. Explain need of NoSQL in detail with examples. [10]
- Q6 Justify the need and benefits of data visualization, explain following tools [10]
- i)matplotlib
- ii) Tableau