

Total No. of Questions : 4]

SEAT No. :

PD138

[6410]-460

[Total No. of Pages : 2

T.E. (I.T.) (Insem)

DATA SCIENCE AND BIG DATA ANALYTICS

(2019 Pattern) (Semester - II) (314452)

Time : 1 Hour]

[Max. Marks : 30

Instructions to the candidates:

- 1) Answer Q.1 or Q.2, Q.3 or Q.4.
- 2) Neat diagrams must be drawn wherever necessary.
- 3) Figures to the right indicate full marks.
- 4) Assume Suitable data, if necessary.

Q1) a) What is Data Explosion? What are key drivers of the data explosion? [8]

b) Write note on Re-engineering in Data Warehouse. [7]

OR

Q2) a) "Big data is having infrastructure challenges" Justify the statement with proper explanation. [8]

b) What are the different Big data learning approaches. Explain with example. [7]

Q3) a) A student recorded her scores on weekly math quizzes that were marked out of a possible 10 points. Her scores were as follows : [8]

8, 5, 8, 5, 7, 6, 7, 7, 5, 7, 5, 5, 6, 6, 9, 8, 9, 7, 9, 9, 6, 8, 6, 6, 7 What is the mean, mode, variance and standard deviation of her scores on the weekly math quizzes?

b) For a cricketer Raj, the chance of getting selected as team caption depends upon the coach of the team. With coach Rahul the probability is 0.5 and with coach Garry the probability is 0.3. Rahul is coach more often about 6 out of 10 tournaments. What is the probability that Raj will be selected as caption for this tournament? [7]

OR

P.T.O.

**Q4) a)** Consider a person uses two brands of bathing soap. Given that person's last purchase of bath-soap was Lux, there is a 90% chance that his next purchase will also be Lux. If he purchased Rexona Soap, there is an 80% chance that his next purchase will also be Rexona. What is the probability that he will purchase Rexona three purchases from now? [8]

b) Write note on [7]

i) Normal Distribution

ii) Variance

iii) Standard Deviation

