

[5353] - 600

T.E. (I.T.) (End-Sem)

DATA SCIENCE AND BIG DATA ANALYTICS

(2015 Pattern)

Time : 2½ hours]

[Max. Marks : 70

Instructions to the candidates:

1) Answer Q1 or Q2, Q3 or Q4, Q5 or Q6 and Q7 or Q8, Q9 or Q10.

2) Neat diagrams must be drawn wherever necessary.

3) Figures to the right side indicate full marks.

4) Assume suitable data if necessary.

Q1) a) What is data warehouse? Explain design and architecture of data warehouse. [6]

b) What is the application of tail and bound in big data. [4]

OR

Q2) a) List and explain data processing infrastructure challenges in Big Data. [6]

b) Explain how Google file system solves big data processing challenges. [4]

Q3) a) Given that a person's last purchase was pepsi, there is a 90% chance that his next purchase will also be pepsi. If a person's last purchase was coke, there is an 80% chance that his next purchases will also be coke. What is the probability that he will purchase pepsi three purchases from new? [6]

b) Explain job execution in Hadoop with example. [4]

OR

Q4) a) Explain the Flajolet Martin Distance sampling. Find the distinct element from the element stream 1, 4, 2, 1, 2, 4, 4, 1, 2, 4, 1, 2, 4, 1, 7. Assume suitable hash function. [6]

b) Prove the principle of linearity of expectation. [4]

P.T.O.

Q5) a) What is data preparation? Explain its types with suitable example. [8]
b) Explain the different modes of data transformation in big data. [8]

OR

Q6) a) What is the need of big data analysis? Explain the different types of analysis techniques. [8]
b) Explain the data analysis life cycle in big data. [8]Q7) a) What are the major challenges in visualizing the big data and how to overcome these challenges. [8]
b) Explain i) Google chat API ii) Cloudera [8]

OR

Q8) a) Explain any two visual data representation techniques with sample data set. [8]
b) Explain i) Data visualization with Tableau ii) Jasper reports [8]Q9) a) What is social media analytics? Explain its need with sample case study. [9]
b) What is text mining? Draw and explain text mining architecture and explain its need. [9]

OR

Q10) a) How mobile analytics is different than social media analytics. Explain with suitable example. [9]
b) Explain roles and responsibilities of big data analyst and data scientist. [9]

❖ ❖ ❖