Total No. of Questions: 8]	96	SEAT No. :
P826	[5870] - 1149	[Total No. of Pages : 2

T. E. (Information Technology) DATA SCIENCE AND BIG DATA ANALYTICS (2019 Pattern) (Semester - II) (314452)

	(201) I atti ii (Semester II) (ST 1182)	
Time :	2½ Hours] [Mo	ax. Marks : 70
	ctions to the candidates:	
1)		
2)		
3) 4)		
7)	Assume the suddie data, if necessary.	
Q1) a	Explain Big data Ecosystem with suitable diagram.	[7]
b	Explain anatomy of File read and write in HDFS.	[7]
c) Write and explain any two Hadoop shell commands.	[4]
	OR O	
Q2) a	Explain Map Reduce with proper diagram for word count	example. [7]
b) Explain Google file system.	[7]
c) Explain ETL processing.	[4]
Q3) a)	Explain different steps in Data Analytics Project Life cycle	[7]
b) Draw and explain Architecture of HIVE.	
c) Explain different data transformation techniques.	[3]
	OR	3
Q4) a)	Explain different kinds of Big Data Analysis.	[7]
b) How data can be ingested in python. Write symax in py	thon for the
	same.	[7]
c) Explain role of visualization in big data analytics.	[3]
Q5) a)	Explain different techniques of Big Data visualization.	[7]
b) Explain challenges in Big data visualization.	[7]
c) Write two data visualization functions from matplotlib.	[3]
	O [™]	

OR

Q6) a)	Explain different tools for data visualization.	
b)	Explain scatter plot, histogram and heat map with example.	
c)	Write two data visualization functions from seaborn.	[3]
Q7) a)	How Social Media analytics helps in value creation? Explain with suit	able
	examples.	[7]
b)	Explain in brief data analytics life cycle.	
c)	Explain big data value terminology.	[4]
	OR	
Q 8) a)	What is text mining? Draw and explain text mining architecture an	d ite
Q0) a)	use.	[7]
b)	Explain Big data analytics in research.	[7]
U)		
c)	Explain big data impact on organizations.	[4]
	And the state of t	
	Ray The Training and the second secon	
	O. Ko.	
[5870] -	2 2	