

Total No. of Questions : 10]

SEAT No. :

P3008

[Total No. of Pages : 2

[5669]-600

T.E. (IT)

DATA SCIENCE AND BIG DATA ANALYTICS

(2015 Pattern)

Time : 2½ Hours]

[Max. Marks : 70

Instructions to the candidates:

- 1) All questions are compulsory.
- 2) Figures to the right indicate full marks.

Q1) a) Explain Big data along with 5V's. [6]

b) Explain pair wise independent hashing. [4]

OR

Q2) a) Explain Machine learning approaches in Big Data. [6]

b) Write two Hadoop shell file management command. [4]

Q3) a) A petrol station owner is considering the effect on his business (superpet) of a new petrol station(Global) which has opened just down the road. Currently (of the total market shared between superpet and Global) Superpet has 80% of the market and Global has 20%. Analysis over the last week has indicated the following probabilities for customers switching the station they stop at each week:

	To		
	Superpet	Global	[6]
From superpet	0.75	0.25	
Global	0.55	0.45	

What will be the expected market share for superpet and Global after another two weeks have past?

b) Explain Map reduce with proper diagram for word count example. [4]

P.T.O.

OR

- Q4)** a) Explain anatomy of file read and write in HDFS. [6]
b) Explain the flajolet martin distance sampling. Find the distinct Element from the element stream 4,2,5,9,1,6,3,7. Consider Hash function $h(x)=(x+6) \bmod 32$ [4]

- Q5)** a) Explain different steps in Data Analytics project life cycle. [8]
b) Explain different kinds of Data Analysis. [8]

OR

- Q6)** a) Draw and explain Architecture of HIVE. [8]
b) Write a user defined function to calculate a cube of given number in R. [8]

- Q7)** a) List the conventional Data visualization tools. Explain any two. [8]
b) Explain pie chart and scatter plot. [8]

OR

- Q8)** a) Describe different types of data visualization. [8]
b) Explain different data visualization tools. [8]

- Q9)** a) Explain Data Scientist role and responsibilities with suitable diagram. [9]
b) How Social Media analytics helps in value creation? Explain with suitable example. [9]

OR

- Q10)** a) What is text mining? Draw and explain text mining architecture and its use. [10]
b) Explain four Big Data use cases. [8]

