NAVRACHANA UNIVERSITY
a UGC recognized University

School: School of Engineering and Technology
Program: BSc(Data Science)
Year: 2nd **Semester:** 4th
Examination: End Semester Examination
Examination year: May – 2023

Course Code: DS209 **Course Name:** Data Analytics using Hadoop
Date: 15-05-2023 **Total Marks:** 40
Time: 10:00 am to 12:00 noon **Total Pages:** 01

Instructions:

➔ Write each answer on a new page.
➔ Use of a simple calculator is permitted.
➔ Draw all relevant waveforms in answer sheet only.
➔ * COs=Course Outcome mapping. # BTL=Bloom's Taxonomy Level mapping

			Marks	CO	BTL
Que. 1.		**Do as directed.**	**[10]**		
	A.	Give two examples of Unstructured data.	[01]	CO1	BT4
	B.	List down V's applicable to BI domain.	[01]	CO1	BT1
	C.	What is the relationship between knowledge, data and information.	[01]	CO1	BT5
	D.	What is wisdom of crowds principle?	[01]	CO1	BT2
	E.	"MORE DATA USUALLY BEATS BETTER ALGORITHMS" Justify.	[02]	CO2	BT3
	F.	What is YARN?	[01]	CO2	BT1
	G.	What is Volunteer Computing?	[01]	CO2	BT1
	H.	What is the difference between skip() & seek() apis.	[01]	CO3	BT2
	I.	The column family and qualifier are separated by character _____	[01]	CO4	BT1
Que. 3.		**Attempt any five.**	**[10]**		
	A.	Explain Namenode in detail.		CO1	BT2
	B.	Explain Datanode in detail.		CO1	BT2
	C.	Explain the concept of Shadow Log.		CO2	BT2
	D.	What do you mean by Operations Log.		CO2	BT2
	E.	Differentiate between RDBMS and HBase		CO5	BT3
	F.	What do you mean by replication factor? What is the default value of replication factor.		CO4	BT2
Que. 3.		**Attempt any two.**	**[10]**		
	A.	Write a short note on the role of Machine Learning in Big Data.		CO1	BT2
	B.	Write a short note on HDFS.		CO3	BT2
	C.	Explain about file permissions on HDFS.		CO3	BT2
Que. 4.		**Attempt any one.**	**[10]**		
	A.	Write a short note on 4V's of Big Data. Explain in brief about different data stored in Big Data.		CO1	BT2
	B.	Write a short note on MapReduce.		CO4	BT2