

K. J. Somaiya Institute of Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

April – May 2024

B.Tech Program: Artificial Intelligence & Data Science Scheme-II

Examination: TY

Semester: VI

Course Code: **AIC604**

Course Name: **Big Data Analytics**

Date of Exam: 25/05/2024

Duration: 2.5 Hours

Max. Marks: 60

Instructions:

- (1) All questions are compulsory.
- (2) Draw neat diagrams wherever applicable.
- (3) Assume suitable data, if necessary.

		Max. Marks	CO	BT level
Q 1	Solve any six questions out of eight.	12		
i)	What is Page Rank? Explain the inverted Index.	02	CO4	R
ii)	What is CURE.	02	CO5	R
iii)	What is HDFS ? What are its features.	02	CO2	R
iv)	Compare Traditional and Big Data.	02	CO1	An
v)	Explain MapReduce framework with suitable example.	02	CO3	U
vi)	What is Bloom filter?	02	CO4	R
vii)	Explain collaborative filtering.	02	CO6	U
viii)	Explain structured, semi structured and unstructured with help of example.	02	CO2	U
Q.2	Solve any four questions out of six.	16		
i)	Explain HITS algorithm.	04	CO6	U
ii)	Explain various tools used in Big Data.	04	CO1	U
iii)	Write Short note on MongoDB..	04	CO2	R
iv)	Explain DGIM algorithm for counting ones in stream with example.	04	CO4	U
v)	Describe Grouping and Aggregation operation with Map Reduce.	04	CO3	U
vi)	Explain PCY algorithm.	04	CO5	U

K. J. Somaiya Institute of Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

April – May 2024

B.Tech Program: Artificial Intelligence & Data Science

Scheme-II

Examination: TY

Semester: VI

Course Code: **AIC604**

Course Name: **Big Data Analytics**

Date of Exam: 25/05/2024

Duration: 2.5 Hours

Max. Marks: 60

Q.3	Solve any two questions out of three.	16		
i)	Explain various application of Big data application.	08	CO1	U
ii)	Explain Flajolet - Martin algorithm in detail.	08	CO4	U
iii)	Explain Spam Farm architecture in details.	08	CO5	U
Q.4	Solve any two questions out of three.	16		
i)	Explain matrix - vector multiplication algorithm by Map Reduce.	08	CO3	U
ii)	Explain different NoSQL data architecture patterns	08	CO2	U
iii)	Explain following clustering algorithm in short: (a) Hierarchical (b)K-means (c) K-mediod (d) Fuzzy C-means	08	CO6	U
