

Total No. of Questions : 5]

SEAT No. :

P570

[Total No. of Pages : 2

[5840]-304

**S.Y. M.Sc. (Computer Science)
CSDT - 234A : BIG DATA ANALYTICS
(2019 Pattern) (Semester - III)**

Time : 2 Hours]

[Max. Marks : 35

Instructions to the candidates:

- 1) *Q. 1 is compulsory.*
- 2) *Solve any 3 question from Q. 2 to Q.5.*
- 3) *Question 2 to 5 carry equal marks.*

Q1) Solve any Five of the following :

[5]

- a) What are the benefits of using Big Data?
- b) Write any two characteristics of Database workload.
- c) Define Master Data Management.
- d) What is hadoop?
- e) Explain real time analytics.
- f) What is Pig?

Q2) Attempt the following :

[10]

- a)
 - i) Explain 3v's of Big Data? **[2]**
 - ii) Define characteristics of Big Data. **[4]**
- b) Write in details step to extract value from Big Data. **[4]**

Q3) Attempt the following :

[10]

- a)
 - i) What is MapReduce? **[2]**
 - ii) Define HBase and Mahout. **[4]**
- b) Explain Hadoop Architecture in detail. **[4]**

P.T.O.

Q4) Attempt the following : **[10]**

- a) i) Explain Hadoop Distributed file system. **[2]**
ii) Write in details application of Big Data. **[4]**
- b) Explain Big Data Analytics Methodology. **[4]**

Q5) Attempt any Two of the following : **[10]**

- a) What are the challenge in Big Data?
- b) Write any five components of data ecosystem?
- c) Explain in details - Data Integration Pattern.

