

Roll No. ....

(193)

9336

Printed Pages—4]

4MCA7/CCE9

**Master of Computer Application (Fourth Semester)**

**(CBCS) Examination, May/June 2019**

**BIG DATA ANALYTICS**

अवधि/Duration : 3 घंटे/Hours]

[पूर्णांक/Max. Marks : 80

[न्यूनतम उत्तीर्णांक/Min. Pass Marks : 32

निर्देश :

1. प्रश्न-पत्र पाँच इकाइयों में विभाजित है । प्रत्येक इकाई में आन्तरिक विकल्प दिया गया है ।
2. प्रत्येक इकाई से एक प्रश्न का उत्तर दीजिए । इस प्रकार कुल पाँच प्रश्नों के उत्तर दीजिए ।
3. सभी प्रश्नों के लिए समान अंक नियत हैं ।
4. जहाँ आवश्यकता हो वहाँ उपयुक्त डाटा माना जा सकता है ।
5. अनुवाद में विसंगति होने पर अंग्रेजी स्वरूप को सही माना जाए ।
6. प्रश्न-पत्र में परीक्षार्थी निर्धारित स्थान पर अपना रोल नम्बर अंकित करें ।

**Instructions :**

1. The Question Paper is divided in five Units. Each unit carries an internal choice.
2. Attempt *one* question from each Unit. Thus attempt *five* questions in all.
3. *All* questions carry equal marks.
4. Assume suitable data wherever necessary.
5. English version should be deemed to be correct in case of any anomaly in translation.
6. Candidate should write his/her Roll Number at the prescribed space on the question paper.

**P.T.O.**

### इकाई I (Unit I)

1. (a) स्मार्ट शहरों के विकास में बिग डेटा विश्लेषण किस प्रकार उपयोगी हो सकता है ? 8  
How big data analysis can be useful in the development of smart cities ?
- (b) बिग डेटा क्या है? इसको वोल्यूम तथा वेलोसिटी के संदर्भ में समझाइये। 8  
What is Big Data ? Discuss it in terms of volume and velocity.

#### अथवा (Or)

2. (a) बिग डेटा के विभिन्न आयामों को समझाइये। 8  
Discuss various dimensions of big data.
- (b) हेल्थकेयर, ट्रांसपोर्टेशन एवं मेडीसिन में बिग डेटा किस प्रकार उपयोगी है ? 8  
Discuss big data in healthcare, transportation and medicine.

### इकाई II (Unit II)

3. (a) बिग डेटा सेट में डेटा बिन्दुओं के बीच समानता अथवा असमानता निकालने के लिये विभिन्न डिस्टेंस निकालने की पद्धतियों का वर्णन कीजिये। 8  
Write a note on different distance measures that can be used to find similarity/dissimilarity between data points in a big data set.
- (b) बिग डेटा में किन्हीं दो सेम्पलिंग टेक्निक को उदाहरण सहित समझाइये। 8  
Describe any *two* sampling techniques for big data with the help of example.

#### अथवा (Or)

4. (a) एक बॉक्स में 15 एप्पल (सेब) हैं, जिनमें से 3 दोषपूर्ण हैं। दो सेब को रैंडमली उठाया जाता है। दोनों में से कोई भी दोषपूर्ण नहीं होगा, इस बात की संभावना (प्रोबेबिलिटी) क्या होगी ? 8  
A box contains 15 apples, 3 of which are defective. Two apples are drawn at random. Find the probability that neither of the two are defective.

- (b) निम्न डेटा सेट का मीन (औसत), बहुलक (मोड) तथा माध्यिका (मीडियन) निकालिये : 8

From the data given below find the mean, mode and median :

आयु (Age)	व्यक्तियों की संख्या (No. of Persons)
18-25	50
26-30	70
31-35	80
36-40	180
41-45	150
46-50	120
51-55	70
56-60	50

### इकाई III (Unit III)

5. (a) हडूप कौन से मोड्स चला सकता है ? 8

What are the modes that a Hadoop can run ?

- (b) हडूप का बिल्डिंग ब्लॉक क्या है ? 8

Discuss brief about building block of Hadoop.

### अथवा (Or)

6. (a) हडूप किस प्रकार स्केल-आउट फीचर को प्रदर्शन को बढ़ाने के लिये उपयोग करता है ? उदाहरण सहित समझाइये। 8

Explain with an example, how Hadoop uses scale-out feature to improve the performance.

- (b) निम्न दो कार्य का Map reduce कार्यान्वित किस प्रकार होगा, स्यूडोकोड की सहायता से समझाइये : 8

Show map reduce implementation for the following *two* tasks using pseudocode :

- (i) दो रिलेशन को Join करना उदाहरण सहित।

Join of two relations with an example

- (ii) एक Map reduce चरण से दो मैट्रिक्स को गुणा करना।

Multiplication of two matrices with one map reduce step.

#### इकाई IV (Unit IV)

7. (a) Map reduce फ्रेमवर्क के लिये API का संक्षिप्त वर्णन कीजिये। 8  
Describe in brief about API for MAP reduce framework.
- (b) Map reduce फ्रेमवर्क में “Shuffle” एवं “Sort” ऑपरेशन को उदाहरण सहित समझाइये। 8  
Describe the operation of “Shuffle” and “Sort” in the MAP reduce framework.  
Explain with the help of an example.

#### अथवा (Or)

8. (a) Map reduce में Avro डेटा सीरीयलाइजेशन तकनीक का वर्णन कीजिये। 8  
Explain Avro Data Serialization technique in Map reduce.
- (b) Map reduce में “Map phase” तथा “Combine phase” का वर्णन कीजिये। 8  
Explain “Map phase” and “Combiner phase” in Map reduce.

#### इकाई V (Unit V)

9. (a) हडूप को इंस्टॉल करते समय कौन-सी कन्फिगरेशन फाइल उपयोग में आती है? Mapred-site.xml का उपयोग क्या है ? 8  
List various configuration files used in Hadoop Installation. What is use of Mapred-site.xml ?
- (b) किसी इनपुट टेक्स्ट फाइल में कितने नंबर हैं, इसकी गिनती करने के लिए Map reduce चरणों को लिखिये। कोड को कम्पाइल एवं रन करने के चरणों को लिखिए। 8  
Write Map reduce step for counting total numbers in the input text file(s). Also write the commands to compile and run the code.

#### अथवा (Or)

10. (a) Mapreduce में “Reducer Phase” को समझाइये। 8  
Explain “Reducer Phase” in Map reduce.
- (b) Map reduce डेटा स्टोरेज से संबंधित समस्याओं का वर्णन कीजिए ? 8  
What are the problems related to Map reduce data storage ?