

Day: Monday  
Date: 13/05/2019

Time: 02.30 PM TO 05.30 PM

Max Marks: 60

S-2019-2803

---

**N.B.:**

- 1) All Questions are **COMPULSORY**
  - 2) Use of non-Programmable calculator is **ALLOWED**
  - 3) Figures to the right indicate **FULL MARKS**
  - 4) Assume Suitable data, if necessary.
- 

**Q.1** What is Big data? Discuss in detail the various characteristics of big data. Give two examples of big data case studies indicating which Vs are satisfied by these case studies? **(10)**

**OR**

**Q.1** Differentiate between data and information. What are the different types of data in big data environment? Illustrate with an example the role of CMS in big data management. **(10)**

**Q.2** Consider the Poker game. After one pair, the next most common hands are two pair and three of a kind. **(10)**

Two Pair: Two cards have one rank, two cards have another rank, and the remaining card has a third rank. (Example: 2♥, 2♠, 5♥, 5♣, K♦)

Three of a Kind: Three cards have one rank and the remaining two cards have two other ranks. (Example: 2♥, 2♠, 2♣, 5♣, K♦)

Calculate the probability of each type of hand, which is more likely?

**OR**

**Q.2** What is Sampling and confidence intervals? Define Bias. Illustrate the various types of biases with example. **(10)**

**Q.3** What do you understand by Linear Regression? Illustrate with an example the various steps and commands involved for performing the linear regression in R. **(10)**

**OR**

**Q.3** What is Predictive Model? Elaborate on the big data analytics lifecycle with neat diagram. **(10)**

**Q.4** Define Virtualization. Elaborate on different types of virtualization technology. Explain the importance and advantage of virtualization technology in big data infrastructure. **(10)**

**OR**

**Q.4** What is NoSQL? Describe in detail about the differences between a traditional relational database and NoSQL. **(10)**

**Q.5** What is Map reduce? Show map reduce implementation for the following two tasks using pseudocode: **(10)**

- a) Multiplication of two matrices, with one map reduce step.
- b) Join of two relations with an example.

**OR**

**Q.5** What is Hadoop? Discuss in detail about Hadoop Distributed File System (HDFS). **(10)**

**Q.6** Explain in detail the collaborative filtering based recommendation system? Describe how it is different from content based recommendation system? **(10)**

**OR**

**Q.6** Differentiate between Privacy and Security? Illustrate in detail with an example how big data is influencing digital marketing aspects? **(10)**