

Roll No

MCA-303 (3)**M.C.A. III Semester (Two Year Course)**

Examination, November 2022

Introduction to Data Science and Big Data

(Elective-I)

Time : Three Hours

Maximum Marks : 70

Note: i) Attempt any five questions.

किन्हीं पाँच प्रश्नों को हल कीजिए।

ii) All questions carry equal marks.

सभी प्रश्नों के समान अंक हैं।

iii) In case of any doubt or dispute the English version question should be treated as final.

किसी भी प्रकार के संदेह अथवा विवाद की स्थिति में अंग्रेजी भाषा के प्रश्न को अंतिम माना जायेगा।

1. a) Write about the modern data analytic tools.
आधुनिक डाटा विश्लेषणात्मक उपकरणों के बारे में लिखिए।
b) Explain Data Visualization.
डाटा विजुअलाइज़ेशन की व्याख्या करें।
2. a) Define Big data. What are the risks of big data? Explain.
बिग डाटा को परिभाषित करें। बिग डाटा के जोखिम क्या हैं? समझाइए।
b) Explain Linear and Logistic regression modeling with an example.
लीनियर और लॉजिस्टिक रिग्रेशन मॉडलिंग को उदाहरण सहित समझाइए।
3. Explain the graphical representation of Univariate, Bivariate and Multivariate analysis in R with an example.
R में Univariate, Bivariate और Multivariate विश्लेषण के आलेखीय निरूपण को उदाहरण सहित समझाइए।

4. a) Write the differences between RDBMS and MongoDB.
RDBMS और MongoDB में अंतर लिखिए।
b) Write about Mongo DB-data base model.
Mongo DB डेटा बेस मॉडल के बारे में लिखिए।
5. a) Write about CRUD operations in detail.
CRUD संचालन के बारे में लिखें।
b) Explain in detail HDFS.
HDFS के बारे में विस्तार से बताइए।
6. Explain the process of data replication in Hive and Spark.
हाइव और स्पार्क में डाटा प्रतिकृति की प्रक्रिया की व्याख्या करें।
7. Explain Stream Data Model and Architecture.
स्ट्रीम डाटा मॉडल और आर्किटेक्चर की व्याख्या करें।
8. Answer any three of the following.
a) Counting oneness in a window
b) Apache Sqoop.
c) Fuzzy modeling.
d) Two way cross tabulation.
e) Statistical inference
निम्नलिखित में से किन्हीं तीन के उत्तर दीजिए।
अ) एक खिड़की में एकता की गिनती
ब) अपाचे स्कूप
स) फजी मॉडलिंग
द) टू वे क्रॉस टेबुलेशन
इ) सांख्यिकीय निष्कर्ष
