



DBX-003-1082002

Seat No. _____

M. Sc. I.T. (Sem. II) (CBCS) (W.E.F. 2016) Examination

July - 2022

CS-08 : Introduction to Big Data & Hadoop

Faculty Code : 003

Subject Code : 1082002

Time : $2\frac{1}{2}$ Hours]

[Total Marks : 70

- 1 (a) Answer the following in **one** sentence : 4
- (1) Puppy linux is one type of Linux flavor. True or False.
 - (2) Give the full form of HDFS.
 - (3) Give the full form of CAP.
 - (4) Give the full form of CRC.
- (b) Answer the following : (any **one**) 2
- (1) Explain virtual box and virtual machine.
 - (2) Explain Name Node, Data Node, and Check point Node.
- (c) Answer the following : (any **one**) 3
- (1) Explain 3V
 - (2) Explain rack awareness and block placement.
- (d) Answer the following : (any **one**) 5
- (1) Explain HDFS Architecture.
 - (2) Explain HDFS and its advance features.
- 2 (a) Answer the following in **one** sentence : 4
- (1) Give the full form of BSP.
 - (2) Give the full form of CLC.
 - (3) What is the symbol of pig prompt ?
 - (4) Pig provides the facilities to create user defined functions. True or False.

- (b) Answer the following : (any **one**) **2**
 (1) Explain slot and container.
 (2) Explain macros.
- (c) Answer the following : (any **one**) **3**
 (1) Explain Load and Dump with example.
 (2) Explain Map Reduce Architecture.
- (d) Answer the following : (any **one**) **5**
 (1) Explain ACID v/s BASE
 (2) Explain Pig Architecture.
- 3** (a) Answer the following in **one** sentence : **4**
 (1) Give the full form of HQL.
 (2) We can execute all DML operations on a view. True or False.
 (3) Give the full form of TT-IDF.
 (4). The data present in the partition can be divided further into _____.
- (b) Answer the following : (any **one**) **2**
 (1) Explain External and Managed Tables.
 (2) Explain Combiner and Compression.
- (c) Answer the following : (any **one**) **3**
 (1) What is Page Rank and Inverted Index.
 (2) Explain Partitions and Buckets.
- (d) Answer the following : (any **one**) **5**
 (1) Explain Hive Architecture.
 (2) Give the difference between Pig and Hive.
- 4** (a) Answer the following in **one** sentence : **4**
 (1) _____ keeps track of where the .META.
 (2) Zoo keeper is open source project ? True or False.
 (3) Observer is similar to _____ in RDBMS.
 (4) Who is between Client and Region Server ?

- (b) Answer the following : (any **one**) **2**
(1) Explain Master and the Region Server.
(2) Explain Sorting.
- (c) Answer the following : (any **one**) **3**
(1) Explain HBase v/s Cassandra.
(2) Explain HBase Coprocessors.
- (d) Answer the following : (any **one**) **5**
(1) Explain HBase Architecture.
(2) Explain Major and Minor compaction.
- 5** (a) Answer the following in **one** sentence : **4**
(1) SSL stands for _____.
(2) _____ is a tool designed to transfer data between Hadoop and relational databases.
(3) EC2 stands for _____.
(4) ETL stands for _____.
- (b) Answer the following : (any **one**) **2**
(1) What is Neo4J ?
(2) How to export data using Sqoop ?
- (c) Answer the following : (any **one**) **3**
(1) What is Storm ?
(2) Explain Impala.
- (d) Answer the following : (any **one**) **5**
(1) Explain Spark Architecture.
(2) What is Flume ? How it works ?
-