

# 315320

**12526**

**3 Hours / 70 Marks**

Seat No. 

--	--	--	--	--	--	--	--

- 
- Instructions* – (1) All Questions are *Compulsory*.
- (2) Answer each next main Question on a new page.
- (3) Illustrate your answers with neat sketches wherever necessary.
- (4) Figures to the right indicate full marks.
- (5) Assume suitable data, if necessary.
- (6) Use of Non-programmable Electronic Pocket Calculator is permissible.
- (7) Mobile Phone, Pager and any other Electronic Communication devices are not permissible in Examination Hall.

**Marks**

- 1. Attempt any FIVE of the following :** **10**
- a) Enlist any two characteristics of Big data.
- b) Write four Hive data types.
- c) Give the significance of Round (double a) and Floor (double a) of Big INT return type.
- d) State the need of apache spark in Big data tools.
- e) List any two benefits of Apache Kafka.
- f) State common plot type of data visualization.
- g) State the use of matplotlib tool in data visualization.

P.T.O.

- 2. Attempt any THREE of the following : 12**
- a) Differentiate big data and traditional data on following terms –
    - i) Volume
    - ii) Velocity
    - iii) Variety
    - iv) Scalability
  - b) Develop a mapreduce program to find max temperature.
  - c) Draw Hive Architecture with its components.
  - d) Enlist and explain four HDFS features.
- 3. Attempt any THREE of the following : 12**
- a) Describe the procedure for Apache pig installation.
  - b) Develop a program to perform word count operation.
  - c) Classify various Big data technologies based on their purpose.
  - d) Draw Hadoop Architecture and Explain the different working components.
- 4. Attempt any THREE of the following : 12**
- a) Develop a program to visualization data using various plots using matplotlib.
  - b) Describe different RDD transformation with examples.
  - c) Explain the key features of MongoDB.
  - d) Describe four NOSQL Data Architecture pattern.
  - e) Classify big data technologies based on data mining and data analytics.

**5. Attempt any TWO of the following :****12**

- a) Explain the following terms –
  - i) Operational big data technology.
  - ii) Analytical big data technologyAlso write four applications of big data with example.
- b) Write the HDFS commands for file operations (Any three with example)
- c) Apply following Hive QL clauses and write code.
  - i) FROM clause
  - ii) WHERE clause
  - iii) ORDER clause

**6. Attempt any TWO of the following :****12**

- a) Draw Kafka cluster Architecture and explain each block in detail.
  - b) Explain following data visualization tools in detail.
    - i) Matplotlib
    - ii) Seaborn
    - iii) Common plot types
  - c) Develop program to create dashboard using plotly tool.
-