

**SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY:: PUTTUR**  
(AUTONOMOUS)

**B.Tech. IV Year I Semester Regular & Supplementary Examinations December-2024**  
**BIG DATA ESSENTIALS**

(CSE with Specialization in Cloud Computing)

**Time: 3 Hours**

**Max. Marks: 60**

(Answer all Five Units 5 x 12 = 60 Marks)

**UNIT-I**

- 1 Distinguish between structure, unstructured and semi-structure data with an examples. CO1 L4 12M

**OR**

- 2 a What is big data analytics? Identify the Classification of Analytics. CO1 L3 6M  
b Illustrate in detail about Hadoop streaming. CO1 L2 6M

**UNIT-II**

- 3 Explain the block, name node and data node in Hadoop file system. CO2 L2 12M

**OR**

- 4 a Describe the dataflow process in Hadoop Distributed file System. CO2 L3 8M  
b Analyze the features of Apache Hadoop. CO2 L4 4M

**UNIT-III**

- 5 a Discuss different types of failures in Classic MapReduce. CO3 L2 6M  
b List out the different types of failures in Yet Another ResourceNegotiator. CO3 L1 6M

**OR**

- 6 Categorize different types of MapReduce input formats. CO4 L4 12M

**UNIT-IV**

- 7 Examine the different execution modes available in Pig. CO4 L3 12M

**OR**

- 8 Examine the Pig Latin Relational Operators. CO2 L4 12M

**UNIT-V**

- 9 Explain with a neat diagram the architecture of Hbase. CO2 L2 12M

**OR**

- 10 Discuss about Hive shell command line interface. CO5 L2 12M

\*\*\* END \*\*\*

54