


SIDDHARTH GROUP OF INSTITUTIONS :: PUTTUR
(Autonomous)

Siddharth Nagar, Narayanavanam Road – 517583

QUESTION BANK (DESCRIPTIVE)
Subject with Code : Big Data (16CS520)

Course & Branch: B.Tech - CSE

Sem: III-B.Tech & I-Sem

Regulation: R16

UNIT –I

- | | |
|--|-----|
| 1. Classification of Digital Data. Explain? | 10M |
| 2. Write about challenges with Big Data? | 10M |
| 3. Explain about 5v's? | 10M |
| 4. What is Big Data Analytics? | 10M |
| 5. What Big Data Analytics isn't? | 10M |
| 6. Explain about classification of Analytics? | 10M |
| 7. (a) Write about top challenges facing big data? | 5M |
| (b) Explain the concept of Basically Available soft state Eventual consistency | 5M |
| 8. Discuss why is big data analytics important? | 10M |
| 9. Write about Data Science? | 10M |
| 10.(a) Explain the difference between parallel and distributed system? | 5 M |
| (b) Explain Cape theorem. | 5M |

UNIT 2

- | | |
|---|-----|
| 1. Compare Reporting and Analysis with its process. | 10M |
| 2. Explain the following | |
| a. Advanced analytics | 3M |
| b. Operationalized analytics | 4M |
| c. Monetized analytics | 3M |
| 3. How to develop an analytical team and what is the skill required for an analyst? | 10M |
| 4. Distinguish statistical significance and business importance. | 10M |
| 5. What are the roles of analytical team and IT team with a detailed note on text analysis? | 10M |
| 6. Explain in detail the commonly used analytical approaches? | 10M |
| 7. Discuss in detail the history of analytical tools. | 10M |

8. How analytical tools have evolved from graphical user interfaces to point solutions to data visualization tools? 10M
9. Give a detailed note on features and limitations of R programming and IBM SPSS. 10M
10. Explain in detail the following
- SAS 5M
 - Compare various analytical tools. 5M

UNIT 3

- (a) List the main feature of Map Reduce. 5M
(b) Explain working of the following phases of Map Reduce with one common example
 - Map Phase 5M
 - Shuffle and sort phase
 - Reducer Phase
- Describe the working of Map reduce with a relevant example. 10M
- Discuss the techniques which is used to optimize the map reduce jobs. 10M
- Discuss the points to be considered while designing a file system in mapreduce. 10M
- What is HBASE? Give detailed note on features of HBASE. 10M
- Write a short note on the Hadoop ecosystem and HDFS architecture. 10M
- How does HDFS ensure data integrity in a Hadoop cluster? 10M
- Discuss the following terms
 - Streaming information access. 3M
 - Low latency information access. 3M
 - Rest and thrift 2M
 - Org.apcahe.hadoop.io.package 2M
- What is Meta data? What information does it provide and explain the role of Name node in a HDFS clusters? 10M
- Define Command line interface using HDFS files and give a brief note on Hadoop-specific file system types and HDFS commands. 10M

UNIT 4

- What is NoSQL? What are the advantages of NoSQL? And Explain types of NoSQL Databases? 10M
- Differentiate between SQL vs NoSQL? 10M

3. What is NewSQL? Differentiate between NewSQL and NoSQL? 10M
4. With Neat sketch explain in detail Hadoop architecture and its components? 10M
5. a) List hadoop distributions 5M
b) Compare Hadoop vs SQL 5M
6. With neat sketch explain HDFS? 10M
7. With neat sketch explain processing data with Hadoop? 10M
8. Explain in detail interacting with Hadoop Ecosystem? 10M
9. List and Explain HDLC commands? 10M
10. What are the limitations of Hadoop 1.0? Explain Hadoop 2: HDFS and Hadoop 2: YARN? 10M

UNIT 5

1. List some key elements of social media. 10M
2. Describe the steps to perform text mining. 10M
3. Discuss some commonly used text mining software. 10M
4. List some common online tools used to perform sentiment analysis. 10M
5. What do you understand by sentiment analysis? 10M
6. Discuss some application areas of mobile analytics. 10M
7. Briefly explain some popular mobile analytics tools available in the market. 10M
8. What is the importance of location –based tracking tools? 10M
9. Discuss the necessity of keeping data secure while conducting analytics. 10M
10. Discuss some fields where mobile analytics can be used. 10M