CCS334-BIG DATA ANALYTICS

Question Bank

UNIT-I

Q.No.	PART-A
1.	Define Big data.
2.	List out the key trends in big data.
3.	Define Web analytics.
4.	Write any five industrial applications of big data.
5.	List out the big data technologies.
6.	Write down any two disadvantages of big data.
7.	Write any four applications of big data.
8.	State the difference between Big data and cloud.
9.	State the difference between inter and trans firewall analytics.
10.	Define unstructured data. Give an example.
Q.No	**PART-B**
1	Explain about industrial example for Big data in deatil.
2	Explain detail about any three big data technologies
3	Write brief notes about Web analytics.
4	Explain detail about Crowd sourcing analytics.
5	Explain about mobile business intelligence with an example
6	Explain detail inter and trans firewall analytics.

Q.No	PART-C
1	Explain detail about any five big data applications with an example.
2.	Explain detail about web analytics in big data.
3.	Explain about industrial example for big data in detail.
4.	Explain about mobile business intelligence with an example.
UNIT-II	
Q.No.	**PART-A**
1.	Define NOSQL.
2.	Define Schemaless databases.
3.	What is Cassandra?
4.	List out the components of Cassandra**.**
5.	Write an advantages of NOSQL.
6.	List out classification of NOSQL database.
7.	Write any two examples of Key-value database.
8.	Write an syntax for materialized view.
9.	Define Graph databases.
10.	List out any four Cassandra example.
11.	Define Cassandra clients.
12.	List out Key Features of NoSQL.

Q.No	PART-B
1	Explain detail about Schema less databases.
2	Explain detail about Cassandra clients with an example.
3	Explain detail about Master-Slave replication.
4	Explain details about materialized views in NOSQL.
5	Explain detail about graph databases with an example.
6	Explain detail about Aggregate Data Model in NOSQL databases.
Q.No	**PART-C**
1	Define Cassandra. Give an example how Cassandra create multiple copies of data in database.
2	Explain detail about graph databases with an real time example.
3	Write brief about following: (i)Master slave replication (ii) Cassandra clients.
	UNIT-III
Q.No.	**PART-A**
1.	Define Map Reduce.
2.	List out failures in classic map reduce.
3.	Define the term MR unit.
4.	List out different types of job schedulers in Hadoop.
5.	List out types of Mapreduce.
6.	List out the features of YARN .
7.	List out the main components of YARN architecture.
8.	Define the term YARN.
9.	Write an advantages and disadvantages of map reduce.
10.	Write an applications of Map Reduce.

Q.No	PART-B
1.	Explain detail about YARN architecture.
2.	Write about failures in classic Map-reduce.
3.	Explain in detail about anatomy of MapReduce job run.
4.	Explain detail about classic Map-reduce.
5.	Discuss about advantages and disadvantages of Mapreduce.
6.	Discuss about application workflow in Hadoop YARN Architecture in detail.

S.No	PART-C
1.	Explain detail about Pig and Pig Latin Scripts with an example.
2.	Explain detail about various job schedulers in job scheduling in YARN.
3.	Write brief about shuffling and sorting in Hadoop Mapreduce.
4.	Write brief about Map Reduce workflows.

UNIT-IV

Q.No.	PART-A
1.	Define the term scaling out.
2.	Define the term name node and data node.
3.	Write down the advantages of Hadoop.
4.	List out types of Hadoop file formats.
5.	Define HDFS.
6.	Define the term Name node and Data node.
7.	Write down the disadvantages of Hadoop file system.
8.	List out types of Hadoop data formats.
9.	Define serialization.
10.	What is Cassandra and its uses?

Q.No	PART-B
1.	Write brief about Hadoop distributed file system.
2.	Explain detail about Hadoop pipes.
3.	Explain detail about serialization in Hadoop.
4.	Explain detail about Avro with an example.
5.	Write brief notes on Cassandra and its functions in big data.
6.	Explain detail HDFS concepts in Hadoop.
Q.No	**PART-C**
1.	Explain detail about Hadoop I/O system.
2.	Explain detail about design of Hadoop distributed file system (HDFS)
3.	Explain detail about File based data structures.
4.	Explain detail about HDFS concepts and Java interface.
	UNIT-V
Q.No.	**PART-A**
1.	Write any four applications of pig grunt in big data.
2.	State the difference between RDBMS and HIVE.
3.	What is meant by Hbase?
4.	What is HiveQL?
5.	Write any four applications of pig grunt in big data.
6.	State the difference between RDBMS and HIVE.
7.	What is meant by Hbase?
8.	What is HiveQL?

Q.No	PART-B
1	Explain detail about Hbase data model in detail.
2	Explain detail about pig data model.
3	Write brief about Pig Latin scripts.
4	Explain detail about HiveQL queries.
5.	Explain detail about Hive data types and file formats.
6.	Explain about Hbase clients with an example.
Q.No	**PART-C**
1.	What is pig? Analyze the pig data model with an example.
2.	Write Brief about following concepts with neat diagram **(a)** Hbase clients **(b)** Praxis
3.	Explain detail about Hbase data model and its implementations with an example.