1. What will be outcome of following command. “pyspark –master local[2]”

a) It will start spark with 2 worker

b) it will start spark with 2 node

c) will give error

d) none of the mentioned

2. Which of the following is true statement?

a) RDD cannot be partitioned

b) RDD can be partitioned

c) RDD cannot be created out of python sequence

d) none

3. Which of the following cannot be a way to create RDD

a) Transform python function into RDD

b) sc.readTextFile()

c) sc.parellelize

4. What we need to do in order to use agg function in spark

a) use groupby

b) orderby

b) having

b) none

5. what is the output of union function of RDD

a) shows data from one RDD

b) Combines Data of both the RDDs

c) no output

d) none

6 – What is big data

a – special kind of data

b – a concept

c – unstructured data set

d – Any kind of data which cannot be processed via usual tools

7.how can we hold RDD for prolong time in RAM

a. by persisting it

b.save on hard-drive

hibernate it

d.none

8. What is spark

a.A tool for big data processing

storage system

C. NOSQL processing engine

D. none of the above

9.What are the operation we can do on RDD

a transformation

b. action

c both A and B

d. only A

10. In which category RDD collect() Method falls.

a. action

b. transformation

c. both

d. none

11. what DAG stands for , in apache spark

a) has no meaning in spark

b) Data analysis control

c) Direct acyclic graph

d) Direct Account graph

12. What is the role of broadcast variable in spark?

a) to keep some data

b) its final variable

c) we can use this for distribute data to all the node in cluster

d) All of the mentioned

13. what is role of accumulation in apache spark

a) it’s a function

b) it’s a method

c) a way to transform RDD

d) Gives us read only variable

14 what does map function does in apache spark

a) maps to given parameters

b) map through all the elements of RDD

c) none

15. What is the role of spark streaming?

a. its processes continuously incoming data

b. it works as a processes

c. a method in spark

d. none

16 Can we run apache spark in standalone mode

a) no

c) yes

17. Which arithmetic operators cannot be used with strings ?

a) +

b) \*

c) –

d) All of the mentioned

18. Which of the following commands will create a list?

a) list1 = list()

b) list1 = [].

c) list1 = list([1, 2, 3])

d) all of the mentioned

19What is the output when we execute list(“hello”)?

a) [‘h’, ‘e’, ‘l’, ‘l’, ‘o’].

b) [‘hello’].

c) [‘llo’].

d) [‘olleh’]

20 Suppose list1 is [1, 5, 9], what is sum(list1) ?

a) 1

b) 9

c) 15

d) Error