**Apache Spark Questions**

1)Find out which one is not action

* sortByKey
* countByKey
* collect()

2)What is best way to set checkpoint interval

* 5-10 multiply by batch interval
* 10 multiply by batch interval

3)Basic RDD operation like find base rdd, find transformation and action in code line.

4)Spark SQL related question like finding output from code.

5)Which is not correct window operation

.Window(Second(3),Second(2))

.Window(Second(3))

.Window(Second(4),Second(2))

5)Spark advantage over Hadoop. Why we use spark over Hadoop, something like that.

6)partitionBy related question. From code we need to find best way to use.

7)K-means code related question. We need to find what should be the return type in line.

8) Graphix : find out output of code.

9)who to write scala code into python, they have given some code. We need to find right one.(Scala code like creation of dataframe from case class row object)

10) accumulator related question, they given code and we need to find what will be the output. Ex. (arry x={1,2,3,4} and some operation like acc=x) options are return 4 or thow exception or (1,2,3,4)