Spark partitions

1. what happens if there are less partitions than executors?   1 point

   Mark only one oval.

   ⚪ Spark will not be able to use all the executors to run in parallel, and the computation result will be correct

   ⚪ Spark will not be able to use all the executors to run in parallel, and the computation result will be wrong

   ⚪ it has no effect

2. what happens if there are more partitions than executors?   1 point

   Mark only one oval.

   ⚪ Spark will not be able to use all the executors to run in parallel

   ⚪ This situation is impossible

   ⚪ All executors will works, and some of them will handle more than one parition

3. Who is responsible to partition the data?   1 point

   Mark only one oval.

   ⚪ First executor

   ⚪ the Driver node

   ⚪ executors voting

   ⚪ The cluster manager

4. what is the criterion to decide if a row is in a certain partition?   1 point

   Mark only one oval.

   ⚪ assigning a row to a partition is random

   ⚪ an object hash is computed for the row, and the value is mapped to one of the parititions

   ⚪ first K rows go to the first partition, the next K to the second etc.

5. can an RDD be partitioned by columns only (vertical partitioning)?   1 point

Mark only one oval.

○ yes

○ no

○ cannot decide based on what we learned so far

6. Is partitioning automatic? Can I influence it?   1 point

Mark only one oval.

○ yes, there is no way to influence it

○ yes, we can influence mapping rows to partitions

○ no. The user has to supply a function that will tell Spark how to map each row to a partition

Google Forms