Programming language options and functional programming

6/6 points (100%)

Quiz, 6 questions

✔ **Congratulations! You passed!**

✔ 1 / 1
points

1.

Which programming languages can be used for using GraphX, the ApacheSpark graph processing engine? (Select all that apply)

☐ R

Un-selected is correct

☐ Scala

Correct
Correct

☐ Python

Un-selected is correct

✔ 1 / 1
points

2.

What is the result of the following code?

```
1   def f(x):
2      return (x+2)*2
3   l = [1,2,3,4]
4   map(f,l)
```

Hint: "map" is the python equivalent to "apply" used in the lecture

This is an example of a correct input format (although others are accepted as well):

[3, 6, 1, 25]

Programming language options and functional programming

[6,8,10,12]

6/6 points (100%)

Quiz, 6 questions

Correct Response
Correct

✓ 1 / 1
points

3.
What is the result of the following python code running on top of ApacheSpark (pyspark) ?

```
1   sc.parallelize([1,2,3,4,5]).reduce(lambda a,b:a+b)
```

15

Correct Response
Correct

✓ 1 / 1
points

4.
What has to be changed in an ApacheSpark program using only functions on the RDD API which has been tested on 1 KB of data if it should run on 1 PB of data?

◉ Nothing

Correct
Correct, ApacheSpark code is already parallel code running on top of the RDD API, it doesn't matter how bit the RDD is

◯ You have to implement it in a parallel way

✓ 1 / 1
points

5.

In which programming languages do you find the most libraries for data science specific tasks?

Programming language options and functional programming

6/6 points (100%)

Quiz, 6 questions

☐ Java

Un-selected is correct

☐ Scala

Un-selected is correct

☐ Python

Correct
Correct

☐ R

Correct
Correct

✔ 1 / 1
     points

6.

Data frames are the central data structure in R and also in python's Pandas. But they are not using ApacheSpark in the background, therefore what is the major limitation of data frames used in R or python's Pandas?

○ They are not as cool as ApacheSpark

◉ They are not running in parallel on multiple machines

Correct

○ They are not OpenSource

○ They are very unstable