



## Spark Lesson 3

[Back to Week 5](#)

**6/6** points earned  
(100%)

Quiz passed!



1 / 1  
points

1.

Check all true statements about the Directed Acyclic Graph Scheduler

---



1 / 1  
points

2.

Why is building a DAG necessary in Spark but not in MapReduce?

---



1 / 1  
points

3.

What are the differences between an action and a transformation? Mark all that apply

---



1 / 1  
points

4.

Generally, which are good stages to mark a RDD for caching in memory?

---



1 / 1  
points

5.

What are good cases for using a broadcast variable? Mark all that apply

---



1 / 1  
points

6.

We would like to count the number of invalid entries in this example dataset:

```
1 invalid = sc.accumulator(0)
2 d = sc.parallelize(["3", "23", "S", "99", "TT"]).foreach(count_invalid)
```

What would be a good implementation of the count\_invalid function?

---

