1. 什么是RDD？
2. 什么是DataFrame/DataSet？
3. 什么是Job/Stage/Task？
4. 什么是persist/cache/checkpoint？