1. 每天的离线数据要处理多长时间
2. 京东：

① Spark Executor具体配置

② 调优之前与调优之后性能的详细对比（例如调整map个数，map个数之前多少、之后多少，有什么提升）

1. 一个项目，从拿到需求到得到报表统计结果，经过哪些实际的工作流程
2. 工作中用到过哪些shell脚本，shell脚本的具体功能
3. 数仓拿到数据之后怎么根据数据特征进行表格的设计，具体怎么维度建模
4. 对于具体的业务要如何设计表格，如何设计日志采集系统
5. 常用的报表工具
6. Spark任务使用什么进行提交，javavEE界面还是脚本
7. 数据仓库每天跑多少张表，全跑还是跑一部分，大概什么时候跑？