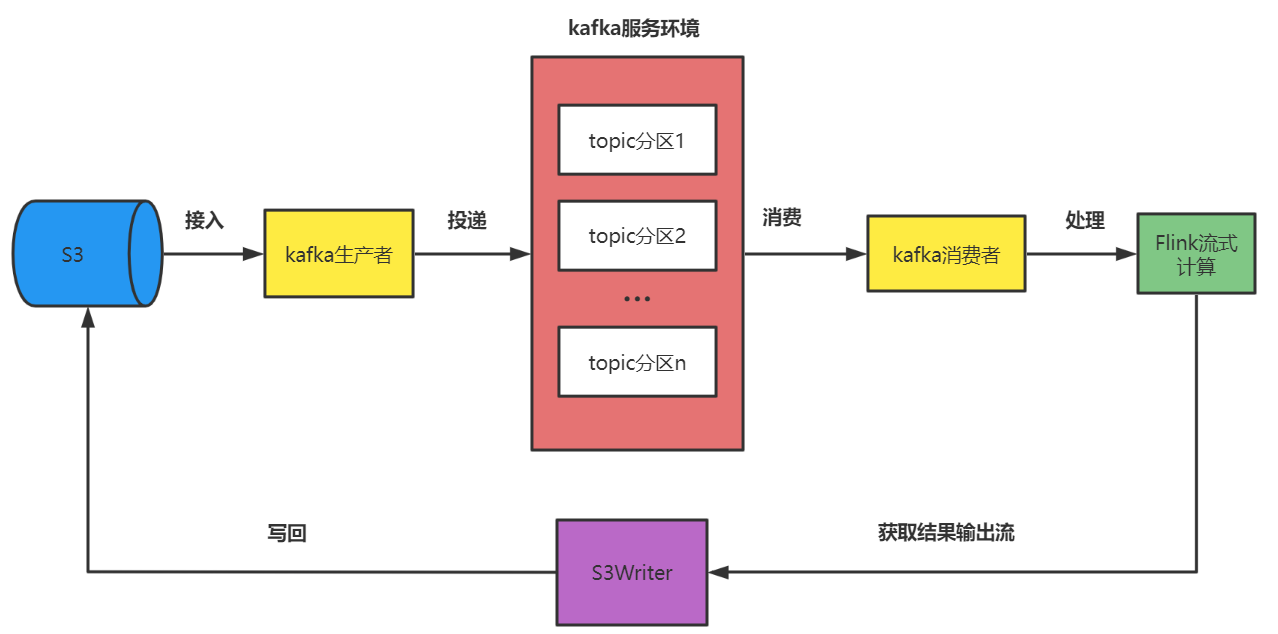
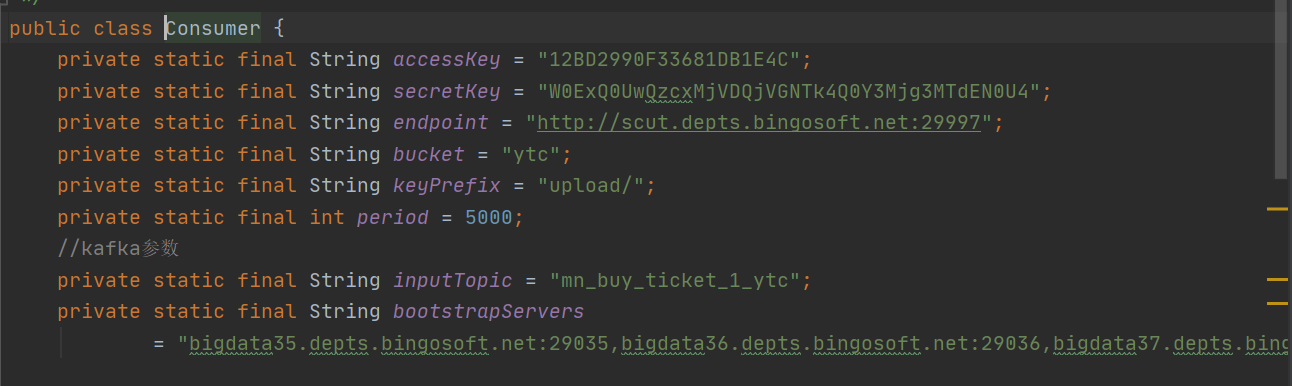
# 流式计算使用说明书

1. **安装环境**
2. jDK 1.8及以上（或JRE 1.8及以上）
3. S3 browser
4. Open VPN
5. IntelliJ IDEA
6. jar依赖包: Flink依赖包 fastjson依赖包
7. **系统流程**

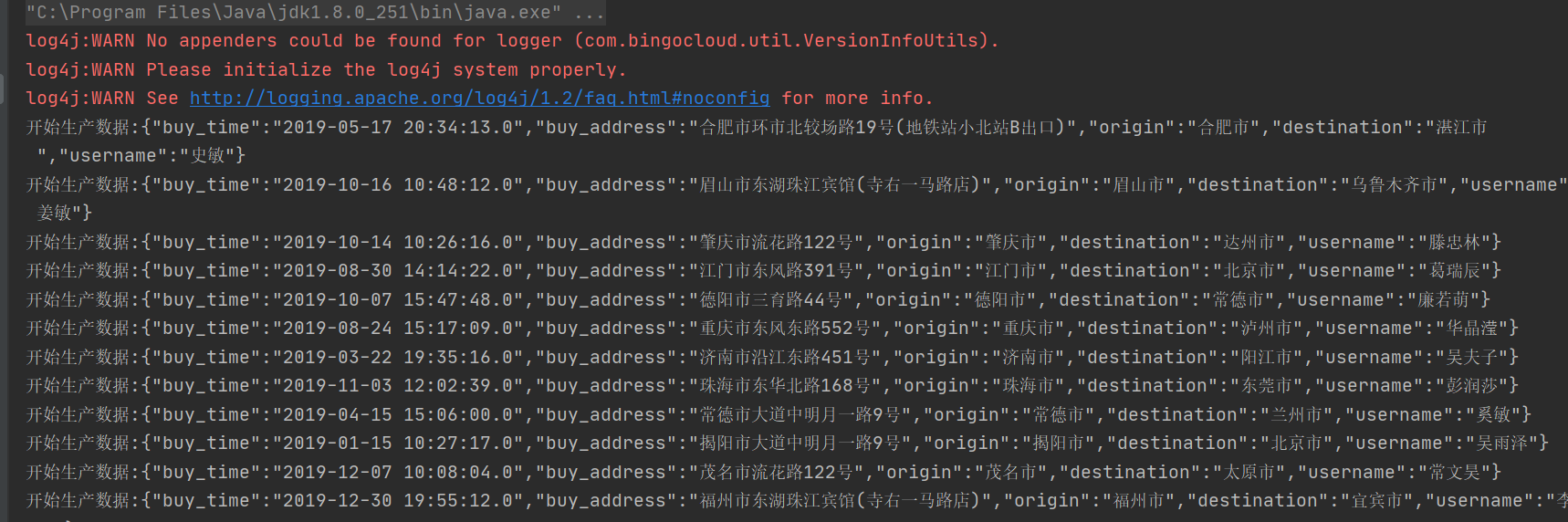


1. **使用步骤**
2. 使用IDEA打开项目文件夹，找到homework文件夹
3. 打开IDEA project structure，在module依赖中添加 Flink依赖和fastjson依赖
4. 在Producer和Consumer类中设置S3连接信息

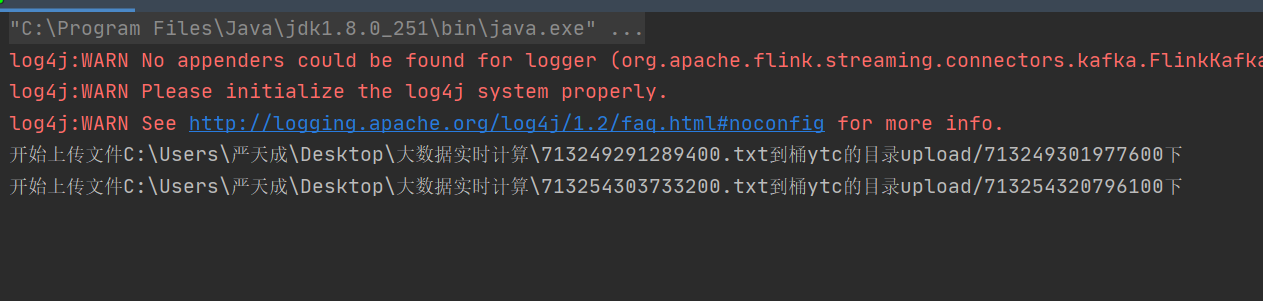


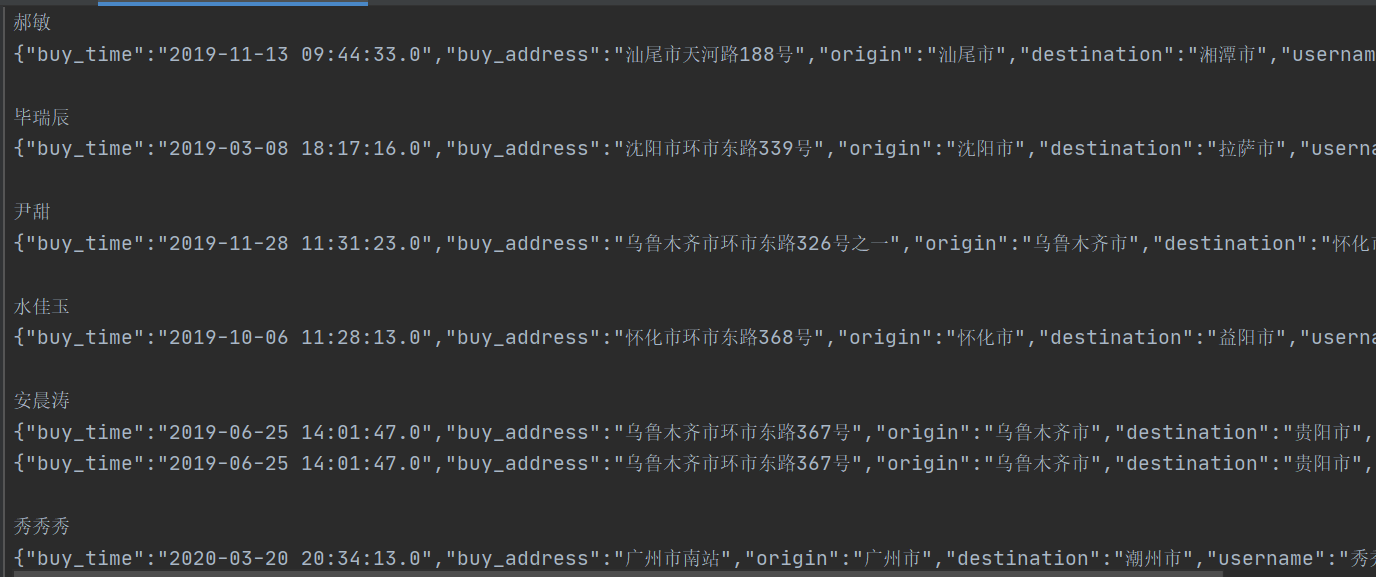


1. 运行Producer的main方法，把S3接入的数据投递到kafka服务环境



1. 运行Consumer的main方法，消费Kafka消息，并交给Flink环境进行流式计算，将不合法的数据过滤（这里比较key的数量），然后username（可在源代码中修改）归类，计算结果上传S3。





1. 查看S3 /upload文件下已存在文件

