

编程题一、将sample.log的数据发送到Kafka中，经过Spark Streaming处理，将数据格式变为以下形式：

commandid | houseid | gathertime | srcip | destip | srcport | destport | domainname | proxytype |
proxyip | proxytype | title | content | url | logid

在发送到kafka的另一个队列中

要求：1、sample.log => 读文件，将数据发送到kafka队列中2、从kafka队列中获取数据（0.10 接口不管理offset），变更数据格式3、处理后的数据在发送到kafka另一个队列中

编程题二、假设机场的数据如下：

1, "SFO"

2, "ORD"

3, "DFW"

机场两两之间的航线及距离如下：

1, 2, 1800

2, 3, 800

3, 1, 1400

用 GraphX 完成以下需求：

- 求所有的顶点
- 求所有的边
- 求所有的triplets
- 求顶点数
- 求边数
- 求机场距离大于1000的有几个，有哪些
- 按所有机场之间的距离排序（降序），输出结果