**Spark SQL/MySQL查询分析器使用说明书**

目录

[1. 启动程序 2](#_Toc42798170)

[1.1前期准备 2](#_Toc42798171)

[1.2 IDEA的配置 2](#_Toc42798172)

[1.3 flink开发环境的搭建 2](#_Toc42798173)

[1.4运行程序 3](#_Toc42798174)

[2. 具体功能 4](#_Toc42798175)

[2.1接入S3的数据 4](#_Toc42798176)

[2.2 flink对接kafka 4](#_Toc42798177)

[2.3 储存至S3中 5](#_Toc42798178)

[3. 其他功能 6](#_Toc42798179)

# 启动程序

## 1.1前期准备

在开始之前，需要做好以下准备：

• 标准的互联网浏览器（建议使用Chrome浏览器）。

• 安装IntelliJ IDEA

• 安装nc.exe（Windows）

• 安装Scala-2.11.8：可参考https://www.runoob.com/scala/scala-install.html

• 下载开发包并解压：

 flink dependencies.zip

具体安装包可通过以下链接进行下载：

http://scuts3.depts.bingosoft.net:29999/resources/default.html

## 1.2 IDEA的配置

按照实操手册要求完成scala 环境的配置。

## 1.3 flink开发环境的搭建

按照实操手册要求完成flink开发环境的搭建，结果如图1-1所示。

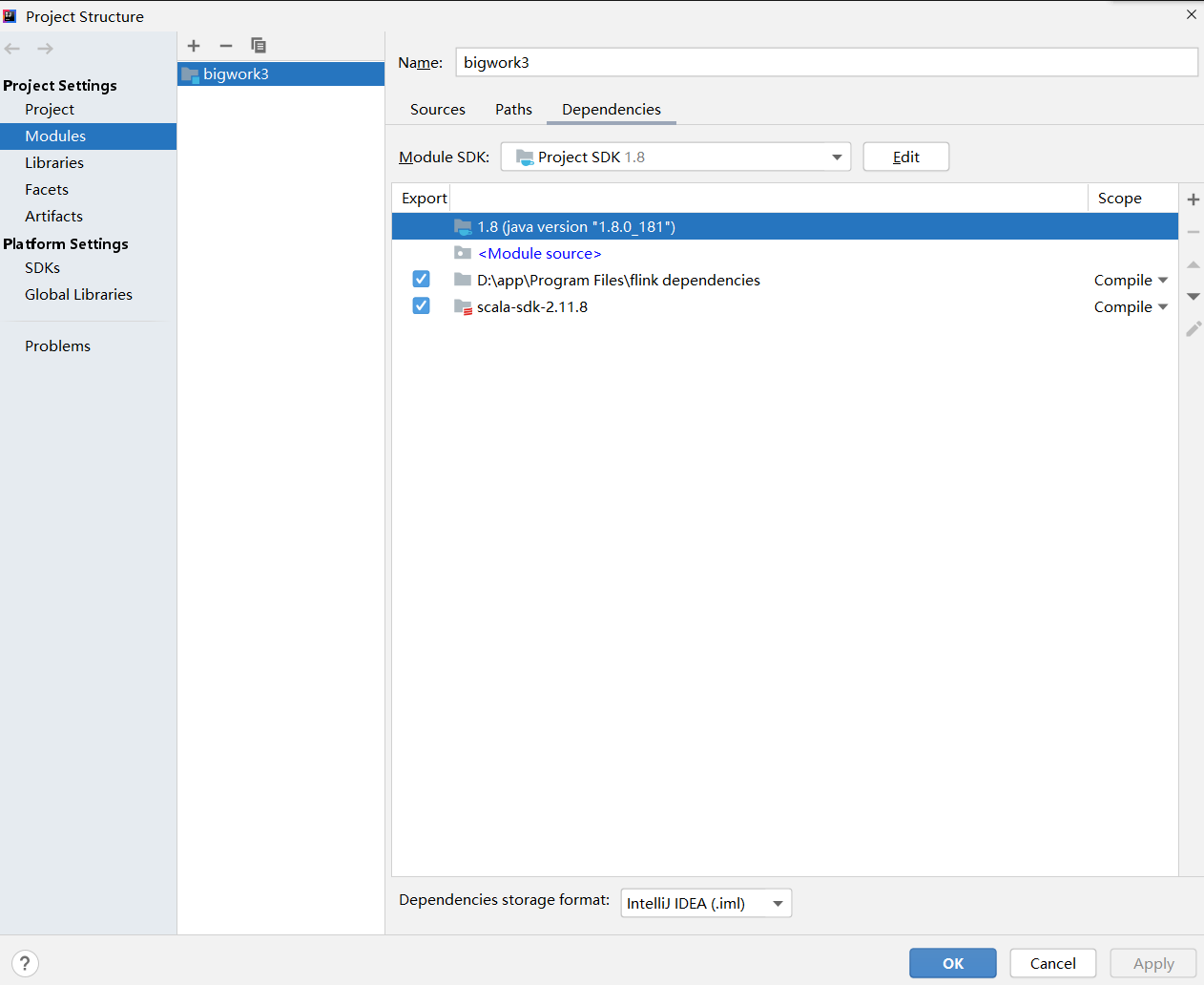


图1-1 完成IDEA的配置以及flink开发环境的搭建

*自此，环境配置工作完成。*

## 1.4运行程序

1.4.1导入源代码至IDEA，如图1-2所示；

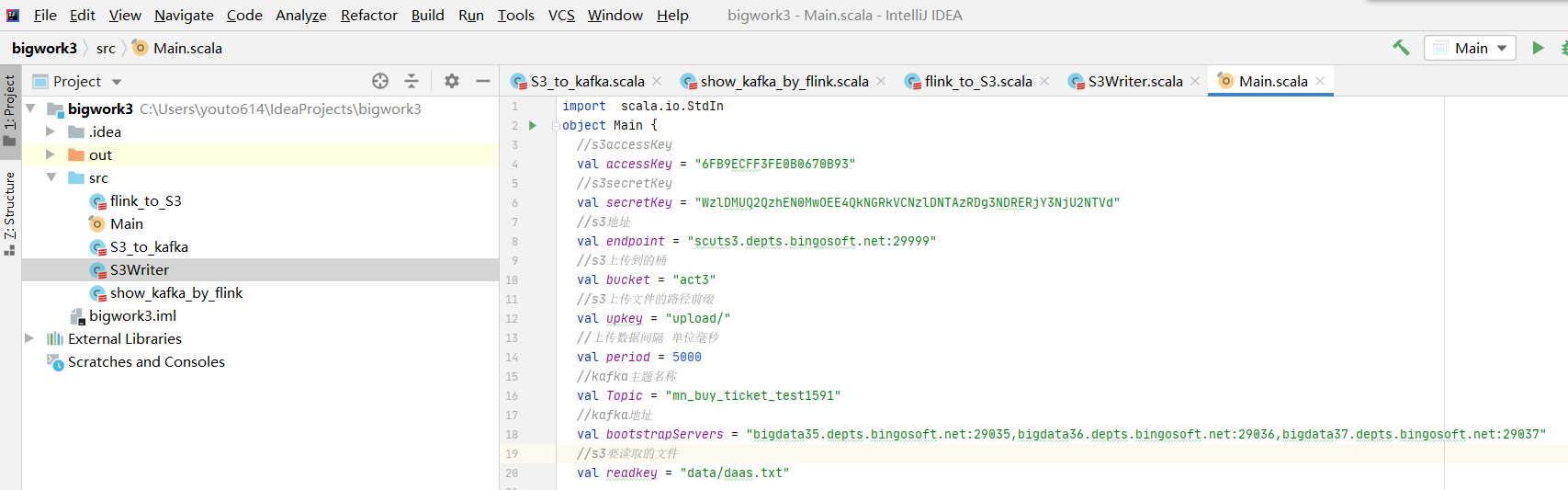


图1-2 导入源码

1.4.2 打开src->Main,在图1-3中可以看到一些具体参数，需要用户根据实际情况进行更改，参数的具体含义已在源码中进行注释。



图1-3 用户需自行修改的参数，图为默认参数

1.4.3 完成参数的修改后，即可点击Run开始运行，运行结果如图1-4所示。

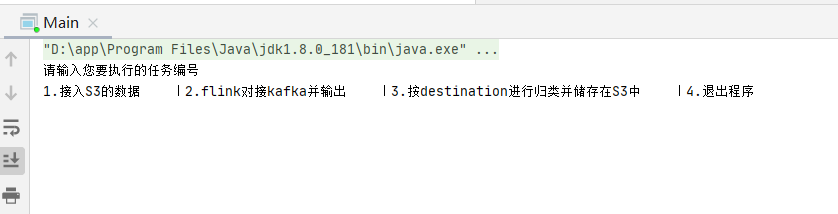


图1-4 运行程序

# 具体功能

## 2.1接入S3的数据

2.1.1在S3由参数指向的桶和路径中，上传相应的文件，如图2-1所示。

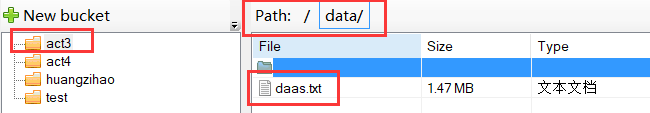


图2-1 图为根据默认参数指定位置上传指定文件

2.1.2 运行程序并在控制台中键入1，即可在控制台中观察到接入S3的数据并生成数据至对应的kafka主题的过程，如图2-2所示；

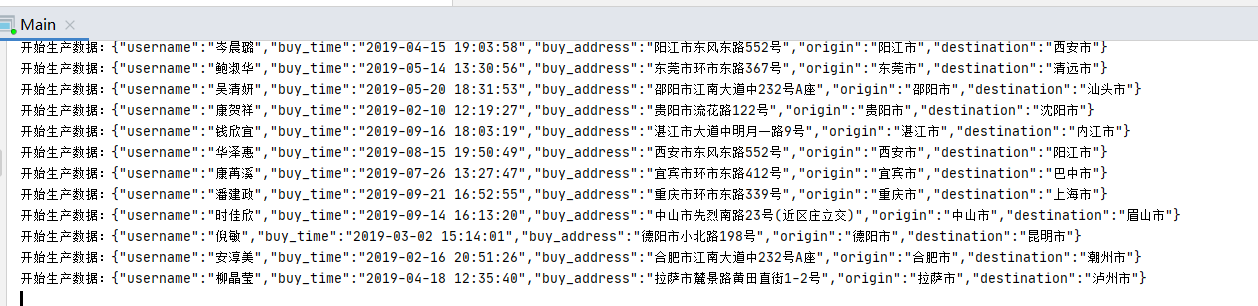


图2-2 生产数据过程

## 2.2 flink对接kafka

运行程序并在控制台中键入2，即可在控制台中观察到flink对接kafk并输出数据的过程，如图2-3所示



图2-3 输出数据过程

## 2.3 储存至S3中

2.3.1运行程序并在控制台中键入3，即可在控制台中观察到按destination进行归类并储存至S3中的过程，如图2-4所示。

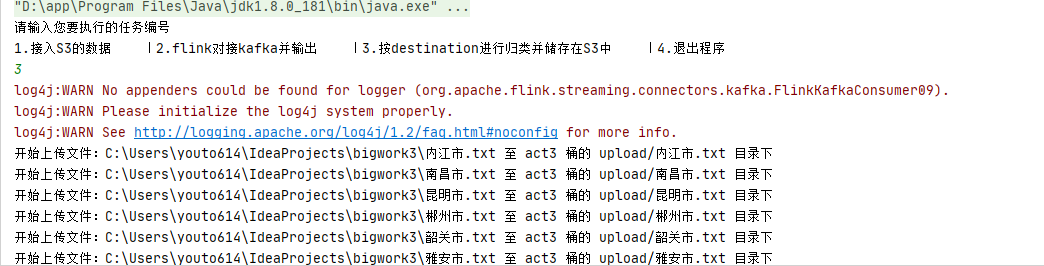


图2-4 归类并储存至S3

2.3.2等待运行完成后，在S3指定位置即可观察到具体文件，如图2-5所示。

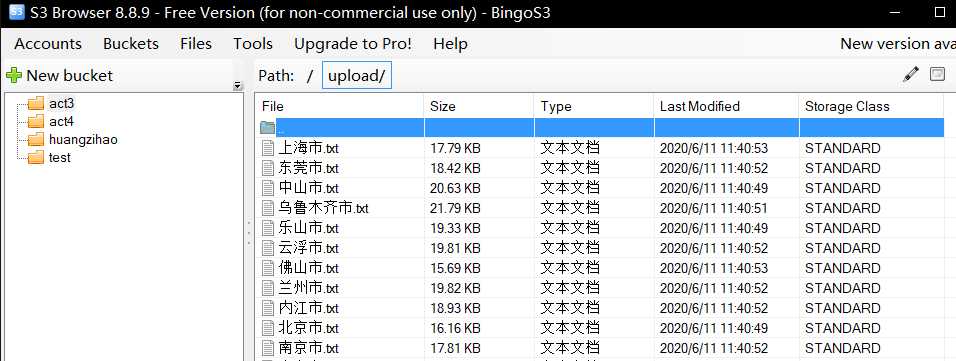


图2-5 文件上传至S3

下载并打开文件，可以看到数据已按destination完成归类，如图2-6所示。



图2-6 数据完成归类

# 其他功能

程序提供处理误输入的功能，用户在误输入的情况下，可以在不关闭程序的情况下重新输入，如图3-1所示。

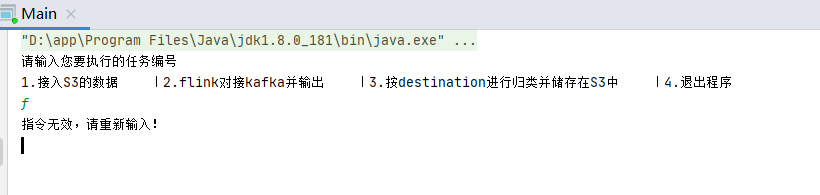


图3-1 处理误输入

程序同时提供关闭功能，保证程序的正常关闭，如图3-2所示。

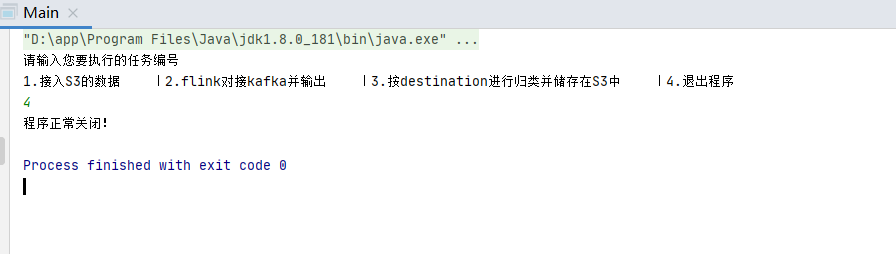


图3-2 程序正常关闭