Glue\_script\_generate模块介绍：

目的：生成一段pyspark代码，也就是glue job里script模块内容（反向生成script,但不会出现图形化界面）

程序主体内容：调用有关接口生成代码段，最后拼接生成完整的glue脚本，并保存到script文件夹

模块内容：source 、transform、target、sql\_query\_parse

第一步sql\_query\_parse：解析sql，传入的sql文件进行解析成source里的表

第二步：开始拼接代码（）

获取head代码：即script里面表头部分

获取source部分代码

获取target代码

获取tali代码: 即script里面尾部部分

第三步：输出py文件到最后路径

需要传入的参数有sql path 和输出文件位置。

流程图：

Ps：如果需要在前台run job，需要在jobdetails配置parms（database、targetpath存放目标结果的s3路径）。或者通过start调glue job