1. 沿袭LinkAnalyzer类

将参数task\_name改为《笑傲江湖》所在的网页URI，headers改为本机浏览器的headers，运行后，实现对《笑傲江湖》页面各章节链接的爬取，并存储在csv文件，置于当前文件夹下。

2.沿袭FileStorageAnalyzer类

实现对《笑傲江湖》各章节html内容的抓取，并存储当前文件夹下的“笑傲江湖”文件夹内。

3.自主创建XAJHAnalyzer类

3.1 parse()方法

3.1.1传入参数

location：html和正文共同存储的文件夹的绝对地址

3.1.2读取

利用os库，依次读取“笑傲江湖”文件夹下的各章节的html文件，

3.1.3解析提取

利用Beautifulsoup进行解析和提取出正文，

3.1.4存储

所提取出的正文以txt形式存储到与html相同文件夹下。

3.2 txt\_cut()方法

3.2.1传入参数：

（1）location：html和正文共同存储的文件夹的绝对地址；

（2）dictionary：《笑傲江湖》人名字典的绝对地址

3.2.2读取人名字典

读取储存有《笑傲江湖》人物名称的字典，并将人名存入列表names

3.2.3读取正文

利用os库，依次读取文件夹下以txt形式储存的各章节正文

3.2.4调用process()方法

对每章特殊字符进行清洗后，调用process()方法，并将process()方法返回的该章人名频数统计结果作为value，储存在字典data\_dict内，并以该章为key

3.2.5调用save()方法

各章遍历完成后，调用save()方法，将统计结果储存在csv文件内，并保存在当前文件夹下。

3.3 process()方法

3.3.1传入参数

（1）dictionary：《笑傲江湖》人名字典的绝对地址

（2）contents：被清洗后的正文内容

（3）names：储存有《笑傲江湖》人名的列表

3.3.2加载字典

加载jieba所需的自定义字典，即参数dictionary

3.2.3统计词频

将人物出现的次数存储在列表counts内，其索引与该任务在列表names中的索引相同

3.2.4返回值

返回列表counts

3.4 save()方法

3.4传入参数

（1）data\_dict：以章节为key，人名频数为value的字典

（2）names：存有《笑傲江湖》人名的列表

3.4.1存储csv

调用pandas库，构造以data\_dict为data、names为index的dataframe，并将dataframe存储到csv文件。