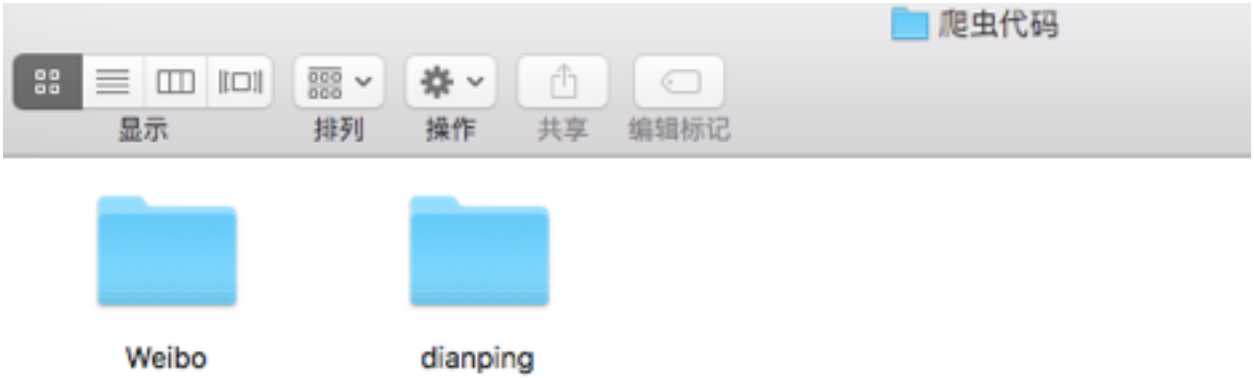
项目配置说明：

一、爬虫模块：

1、文件结构：包含Weibo和dianping两个爬虫文件，分别对应微博数据和大众点评爬虫。



2、安装Python2.7：从官网上下载Python2.7。官方网址：https://www.python.org/download

3、安装Python包管理工具pip。

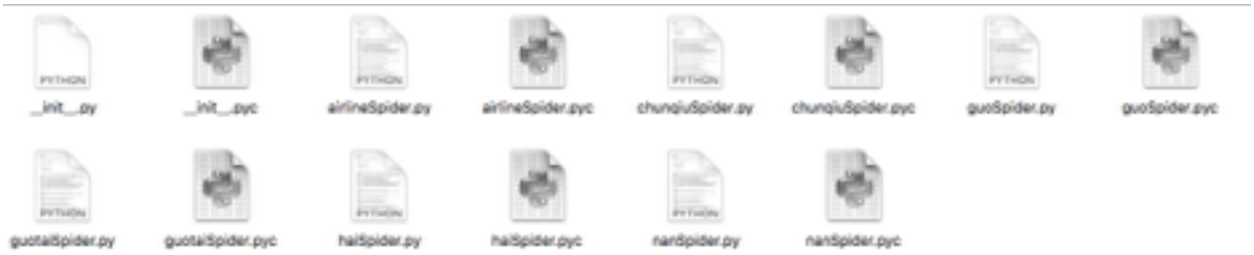
从网址：https://bootstrap.pypa.io/get-pip.py 上得到get-pip.py文件，使用命令安装pip：

```
[~] python get-pip.py
```

4、使用pip命令安装Scrapy。

```
[~] pip install Scrapy
```

5、Weibo文件中有spiders文件夹，其中包含着所有的爬虫。



6、使用命令进入Weibo文件夹中，并通过scrapy crawl +爬虫名爬取数据。其中chunqiuSpider为爬虫名。

```
[Spider] cd Weibo                                                    16:5
[Weibo] scrapy crawl chunqiuSpider                                   16:5
2017-06-18 17:03:06 [scrapy] INFO: Scrapy 1.2.0 started
2017-06-18 17:03:06 [scrapy] INFO: Overridden settings:
torial.spiders', 'ROBOTSTXT_OBEY': True, 'SPIDER_MODULES
 'BOT_NAME': 'tutorial', 'COOKIES_ENABLED': False, 'DEFA
al.items.DmozItem', 'DOWNLOAD_DELAY': 1}
```

二、数据分析模块：

1、使用pip安装ipython和notebook

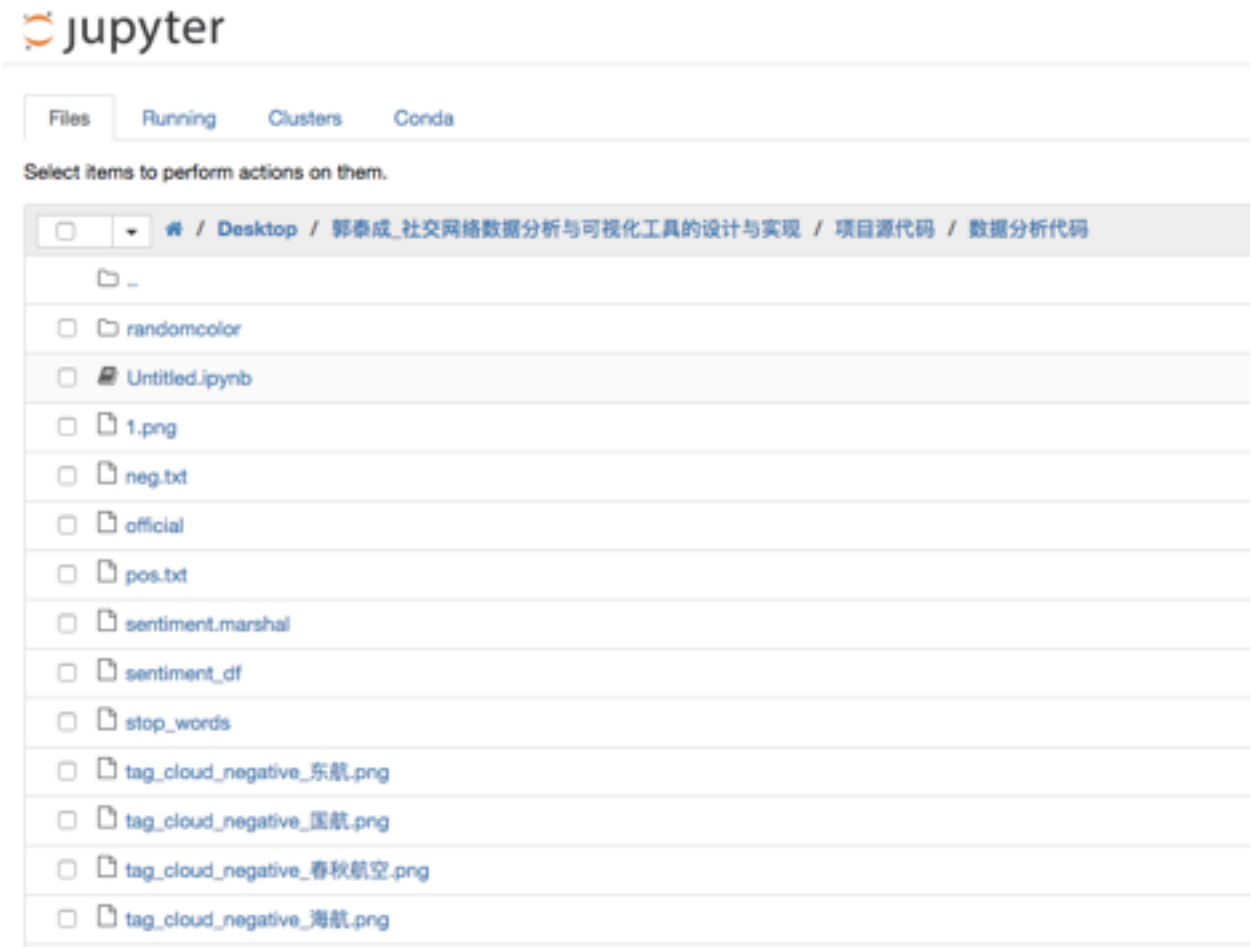
```
[~] pip install ipython
```

```
[~] pip install notebook
```

2、开启交互式IPython

```
[~] ipython notebook
```

3、打开数据分析源代码中的Untitled.ipynb即可运行项目。



项目使用说明：

针对每一条语句，点击图中红色标记的按钮或者按住shift+enter即可运行。