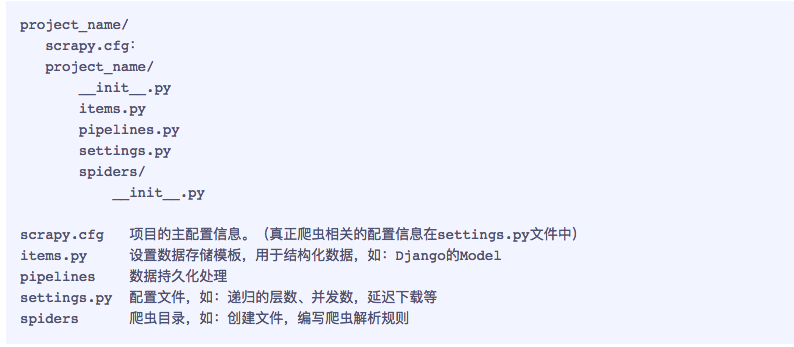
Scrapy框架课程介绍：

1. 框架的简介和基础使用
2. 持久化存储
3. 代理和cookie
4. 日志等级和请求传参
5. CrawlSpider
6. 基于redis的分布式爬虫
7. scrapy框架的简介和基础使用
   1. 概念：为了爬取网站数据而编写的一款应用框架，出名，强大。所谓的框架其实就是一个集成了相应的功能且具有很强通用性的项目模板。（高性能的异步下载，解析，持久化……）
   2. 安装：
      1. linux mac os:pip install scrapy
      2. win:
         1. pip install wheel
         2. 下载twisted：<https://www.lfd.uci.edu/~gohlke/pythonlibs/#twisted>

pip install 下载好的框架.whl

* + - 1. pip install pywin32
      2. pip install scrapy
  1. 基础使用： 使用流程
     1. 创建一个工程：scrapy startproject 工程名称
        1. 目录结构：
     2. 在工程目录下创建一个爬虫文件:
        1. cd 工程
        2. scrapy genspider 爬虫文件的名称 起始url
     3. 对应的文件中编写爬虫程序来完成爬虫的相关操作
     4. 配置文件的编写（settings）
     5. 执行
     6. 在工程目录下创建一个爬虫文件
        1. cd 工程
        2. scrapy genspider 爬虫文件的名称 起始的url
     7. 对应的文件中编写爬虫程序来完成爬虫的相关操作
     8. 配置文件的编写（settings）
        1. 19行：对请求载体的身份进行伪装
        2. 22行：不遵从robots协议
     9. 执行 ：scrapy crawl 爬虫文件的名称 --nolog(阻止日志信息的输出)