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

代理和cookie：

cookie:豆瓣网个人登录，获取该用户个人主页这个二级页面的页面数据。

如何发起post请求？

一定要对start\_requests方法进行重写。

1. Request（）方法中给method属性赋值成post
2. FormRequest（）进行post请求的发送

代理：

下载中间件作用：拦截请求，可以将请求的ip进行更换。

流程：

1. 下载中间件类的自制定
   1. object
   2. 重写process\_request(self,request,spider)的方法
2. 配置文件中进行下载中间价的开启。

