1. 使用Scrapy Shell的控制台，和Scrapy进行交流，可用于测试Scrapy代码语句，而不需要写一个整体的爬虫程序。

例如: Scrapy shell 网页url地址，则进入shell后，response对象会对被加载成所指定的 URL地址。

1. 浏览器可以自动生成选定节点的Xpath。
2. Response.xpath(‘/html/’).extract() 用于查看下载页面的所有内容。可以查看是否成功的下载页面，而没有被拒绝403访问等等。
3. 服务器端为了反爬虫，会对请求的头部进行验证。根据在服务器端设置的验证不同，对请求头部的验证也会不同，有些会验证User-Agent, 有些会验证host. 如果这请求中，这些条目没有设置或者设置错误，则服务器端会拒绝该请求访问。