创建项目：scrapy startproject xxx

创建爬虫：scrapy genspider -t crawl 爬虫名 域名

反爬虫策略：

* 动态设置User-Agent( 随机切换User-Agent，以模拟不同用户的浏览器信息)
* 禁用Cookies（也就是不启用cookies middleware, 不向Server发送cookies ，因为有些网站通过cookies的使用发现爬虫行为）
* Scrapy可以通过settings.py 里的COOKIES\_ENABLED 控制CookiesMiddleware开 启或关闭
* 设置延迟下载

DOWNLOAD\_DELAY = 3

* 使用IP地址池：VPN和代理IP