0、本程序用于爬取京东中文商品评论，字段包含 **商品名称、商家名称、价格、整体好评度、总评论数、晒图、追评、好评、中评、差评、评论者ID、评论、星级、赞、回复数等，**见jd\_jiadian-sample.xml。

1、建数据库表sql1.txt --在Navicat软件中对Mysql数据库建表。

2、获取缓存cookies.txt --用于身份验证，浏览器登录京东主页，按F12复制Cookies。



3、运行ipp.py --用于**更新代理IP**，运行前从以下网站获取http高匿代理IP，存储至ips.txt，ips-1.txt用于存储更新后IP。

https://ip.ihuan.me/tqdl.html

http://www.thebigproxylist.com/

https://topsiteproxy.com/

https://superproxylist.com/

4、运行jd2.py --进行多线程**多进程**爬虫。参数说明：

输入参数：进程数-电脑核数\*2，每类页数-30个商品/页

商品参数：本程序通过检索形式获取商品列表，可用于按类爬（如音乐）或按商品爬（如小米10），在urls中放入爬取对象即可，从谷歌浏览器得到的keyword会以%E6\*\*开头。