

id	name	author	intro	date	price	Press	point
1	追风筝的人 [美] 卡勒德·[为你, 千			2006-5	29.00元	上海人民出	8.9
2	解忧杂货店 [日] 东野圭[一碗精心			2014-5	39.50元	南海出版	8.5
3	小王子 [法] 圣埃克[献给长成			2003-8	22.00元	人民文学出	9.0
4	白夜行 [日] 东野圭[暗夜独行			2008-9	29.80元	南海出版	9.1
5	围城 钱锺书 [对于“人			1991-2	19.00	人民文学出	8.9
6	三体 刘慈欣 [你我不过			2008-1	23.00	重庆出版社	8.8
7	嫌疑人X的献身 [日] 东野圭[数学好是			2008-9	28.00	南海出版	8.9
8	活着 余华 [生的苦难			2012-8-1	20.00元	作家出版社	9.3
9	挪威的森林 [日] 村上春[村上之发			2001-2	18.80元	上海译文出	8.0
10	百年孤独 [哥伦比亚] 加[尼采所谓			2011-6	39.50元	南海出版	9.2
11	红楼梦 [清] 曹雪芹 [谁解其中			1996-12	59.70元	人民文学出	9.6
12	看见 柴静 [在这里看			2013-1-1	39.80元	广西师范大	8.8
13	平凡的世界 路遥 [中国当代			2005-1	64.00元	人民文学出	9.0
14	三体II 刘慈欣 [无边的黑			2008-5	32.00	重庆出版社	9.3
15	不能承受的 [捷克] 米兰[朝向媚俗			2003-7	23.00元	上海译文出	8.5
16	三体III 刘慈欣 [终章, 何			2010-11	38.00元	重庆出版社	9.2
17	达·芬奇密码 [美] 丹·布朗 [一切畅销			2004-2	28.00元	上海人民出	8.2
18	我们仨 杨绛 [平淡, 坚			2003-7	18.80元	生活·读书	8.7
19	天才在左 疯子在右 [简称“疯子			2010-2	29.80元	武汉大学	8.1

```

douban.py x jingdong.py tianqiyubao.py
1 from lxml import etree
2 from random import random
3 import requests
4 import time
5 import pymysql
6
7 conn = pymysql.connect(host='localhost',user='root',password='123456',db='mjl',port=3306,charset='utf8')
8 cursor=conn.cursor()
9 headers={'User-Agent':'Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/63.0.3239.26 Safari/537.
10
11 def get_movies(url):
12     proxies = {
13         "http": "http://119.102.129.114:9999",
14         "https": "http://119.102.129.114:9999",
15     }
16     res=requests.get(url,headers=headers, proxies=proxies)
17     if res.status_code==200:
18         selector=etree.HTML(res.text)
19         infos=selector.xpath('//tr[@class="item"]')
20         for info in infos:
21             name=info.xpath('td/div/a/@title')[0]
22             url=info.xpath('td/div/a/@href')[0]
23             book_infos=info.xpath('td/p/text()')[0]
24             author=book_infos.split('/')[0]
25             Press=book_infos.split('/')[1]
26             date=book_infos.split('/')[2]
27             price=book_infos.split('/')[3]
28             print(info.xpath('td/div/a/@title')[0],book_infos.split('/')[1],book_infos.split('/')[2],book_infos.split('/')[3])

```

```

1 import requests
2 from lxml import etree
3 import time
4 import MySQLdb
5
6 conn=MySQLdb.connect(host='localhost',user='root',passwd='123456',db='mjl',charset='utf8')
7 cur=conn.cursor()
8 headers = {'authority': 'search.jd.com',
9            'method': 'GET',
10            'scheme': 'https',
11            'referer': 'https://search.jd.com/Search?keyword=%E6%89%8B%E6%9C%BA&enc=utf-8&qst=1&rt=1&stop=1&vt=2&wq=%E6%89%8B%E6%9C%BA&enc=utf-8&qst=1&rt=1&stop=1&vt=2&cid2=653&cid3=655&page=' + str(2)*n
12            'user-agent': 'Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/72.0.3626.119 Safari/537.36',
13            'x-requested-with': 'XMLHttpRequest',
14            'cookie': 'qpsc=3; pinId=RAGa4XM0vrs; xtest=1210.cf6b6759; ipLocation=%u5E7F%u4E1C; _jrda=5; TrackID=1aUdbc9HHS2MDEzabuy'
15        }
16
17 def get_mobile(n):
18     url='https://search.jd.com/Search?keyword=%E6%89%8B%E6%9C%BA&enc=utf-8&qst=1&rt=1&stop=1&vt=2&cid2=653&cid3=655&page=' + str(2)*n
19     res=requests.get(url, headers=headers)
20     res.encoding='utf-8'
21     htmltext=etree.HTML(res.text)
22     datas=htmltext.xpath('//li[contains(@class,"gl-item")]')
23     mobile_list=[]
24     for data in datas:
25         array_price=data.xpath('./div/div[3]/strong/i/text()[1]')
26         price=" ".join(map(str, array_price))
27         array_intro=data.xpath('./div/div[4]/a/i/text()')
28         intro=" ".join(map(str, array_intro))

```

```
douban.py x jingdong.py x tianqiyubao.py x
1 from lxml import etree
2 import requests
3 import time
4 import pymysql
5 from random import random
6 from bs4 import BeautifulSoup
7 from bs4 import UnicodeDammit
8 import urllib.request
9
10 conn = pymysql.connect(host='localhost',user='root',password='123456',db='mjl',port=3306,charset='utf8')
11 cursor=conn.cursor()
12 headers={'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/63.0.3239.26 Safari/537.36'}
13
14 def get_weather(url):
15     proxies={"http": "http://119.102.129.114:9999",
16             "https": "http://119.102.129.114:9999"}
17     if url=="http://www.weather.com.cn/weather/101010100.shtml":
18         city="北京"
19     elif url=="http://www.weather.com.cn/weather/101020100.shtml":
20         city="上海"
21     elif url=="http://www.weather.com.cn/weather/101280601.shtml":
22         city="广州"
23     elif url=="http://www.weather.com.cn/weather/101280601.shtml":
24         city="温州"
25     req=urllib.request.Request(url,headers=headers)
26     data=urllib.request.urlopen(req)
27     data=data.read()
28     dammit=UnicodeDammit(data ["utf-8" "gbk"])
```

[illegible]

<div><div>☰</div><div>开始事务</div><div>备注</div><div>筛选</div><div>排序</div><div>导入</div><div>导出</div></div>				
id	city	date	weather	temp
▶ 1	北京	26日 (今天)	多云	10°C
2	北京	27日 (明天)	小雨转多云	16°C/6°C
3	北京	28日 (后天)	多云	20°C/9°C
4	北京	29日 (周一)	多云转小雨	24°C/13°C
5	北京	30日 (周二)	多云	26°C/14°C
6	北京	1日 (周三)	多云	25°C/14°C
7	北京	2日 (周四)	晴	28°C/15°C
8	上海	26日 (今天)	晴	11°C
9	上海	27日 (明天)	多云转小雨	18°C/14°C
10	上海	28日 (后天)	小雨	22°C/17°C
11	上海	29日 (周一)	中雨	22°C/17°C
12	上海	30日 (周二)	阴转多云	20°C/17°C
13	上海	1日 (周三)	多云转阴	25°C/17°C
14	上海	2日 (周四)	阴转多云	23°C/11°C
15	广州	26日 (今天)	中雨	24°C
16	广州	27日 (明天)	中雨转雷阵雨	28°C/24°C
17	广州	28日 (后天)	雷阵雨	29°C/24°C
18	广州	29日 (周一)	雷阵雨转中雨	30°C/25°C
19	广州	30日 (周二)	中雨转大雨	29°C/23°C
20	广州	1日 (周三)	大雨转多云	26°C/22°C
21	广州	2日 (周四)	多云	28°C/21°C
22	温州	26日 (今天)	阴	15°C
23	温州	27日 (明天)	阴转小雨	20°C/16°C

豆瓣图书TOP250 · 京东商城 · 天气预报			
<div>书名：追风筝的人</div> <div>作者：[美] 卡勒德·胡赛尼</div> <div>出版社：上海人民出版社</div> <div>出版日期：2006-5</div> <div>价格：29.00元</div> <div>评分：8.9</div> <div>内容：['为你，千千万万遍']</div>	<div>书名：解忧杂货店</div> <div>作者：[日] 东野圭吾</div> <div>出版社：南海出版公司</div> <div>出版日期：2014-5</div> <div>价格：39.50元</div> <div>评分：8.5</div> <div>内容：['一碗精心熬制的东野牌鸡汤，拒绝很难']</div>	<div>书名：小王子</div> <div>作者：[法] 圣埃克苏佩里</div> <div>出版社：人民文学出版社</div> <div>出版日期：2003-8</div> <div>价格：22.00元</div> <div>评分：9.0</div> <div>内容：['献给长成了大人的孩子们']</div>	<div>书名：白夜行</div> <div>作者：[日] 东野圭吾</div> <div>出版社：南海出版公司</div> <div>出版日期：2008-9</div> <div>价格：29.80元</div> <div>评分：9.1</div> <div>内容：['暗夜独行的残破灵魂，爱与恶本就难分难舍']</div>
<div>书名：围城</div> <div>作者：钱锺书</div> <div>出版社：人民文学出版社</div> <div>出版日期：1991-2</div> <div>价格：19.00</div> <div>评分：8.9</div> <div>内容：['对于“人艰不拆”四个字最彻底的违抗']</div>	<div>书名：三体</div> <div>作者：刘慈欣</div> <div>出版社：重庆出版社</div> <div>出版日期：2008-1</div> <div>价格：23.00</div> <div>评分：8.8</div> <div>内容：['你我不过都是虫子']</div>	<div>书名：嫌疑人X的献身</div> <div>作者：[日] 东野圭吾</div> <div>出版社：南海出版公司</div> <div>出版日期：2008-9</div> <div>价格：28.00</div> <div>评分：8.9</div> <div>内容：['数学好是一种极致的浪漫']</div>	<div>书名：活着</div> <div>作者：余华</div> <div>出版社：作家出版社</div> <div>出版日期：2012-8-1</div> <div>价格：20.00元</div> <div>评分：9.3</div> <div>内容：['生的苦难与伟大']</div>
<div>书名：挪威的森林</div> <div>作者：[日] 村上春树</div> <div>出版社：上海译文出版社</div> <div></div> <div></div> <div></div> <div></div>	<div>书名：百年孤独</div> <div>作者：[哥伦比亚] 加西亚·马尔克斯</div> <div>出版社：南海出版公司</div> <div></div> <div></div> <div></div> <div></div>	<div>书名：红楼梦</div> <div>作者：[清] 曹雪芹 著</div> <div>出版社：人民文学出版社</div> <div></div> <div></div> <div></div> <div></div>	<div>书名：看见</div> <div>作者：梁静</div> <div>出版社：广西师范大学出版社</div> <div></div> <div></div> <div></div> <div></div>

豆瓣图书TOP250

京东商品

天气预警

产品名称：【预售】魅族 16s 骁龙855全面屏拍照游戏，6GB+128GB 碳纤维 全网通移动联通电信4G，双卡双待

价格：3198.00

简介：

产品名称：Apple iPhone XR (A2108) 128GB 黑色 移动联通电信4G，双卡双待

价格：5698.00

简介：【4.25Apple新品日未裁，抢券再减400元！】高性价比iPhone8仅售3499元，小屏经典iPhone7限量特价！更有大眼优惠部，白条券包等你来抢！

产品名称：【KPL官方比赛用机】vivo iQOO 44G超快闪充 8GB+128GB电竞蓝 全面屏拍照，骁龙855电竞游戏 全网通4G

价格：3298.00

简介：【白条3期先电-赠蓝牙耳机！赠单件立减50元E卡！】高通骁龙855，44G超快闪充，4000mAh大电池，非凡电竞体验X27新品上市，立即抢购

产品名称：荣耀8X 千元屏霸 91%屏占比 2000万AI双摄 4GB+64GB 幻夜黑 移动联通电信4G全面屏，双卡双待

价格：1299.00

简介：限时优惠1299！麒麟710处理器，2000万AI双摄，人脸+指纹双识别！荣耀瀑布屏特惠，送品质，购荣耀~

产品名称：小米 Redmi Note7 幻影渐变AI双摄 4GB+64GB 梦幻蓝 全网通4G 双卡双待 水滴全面屏拍照游戏智能

价格：1199.00

简介：【品质特惠】4800万拍照千元机，品质“小金刚”，18个月质保，”满血版“骁龙660，4000mAh大电量！品质好物选小米，京东选购更省钱

豆瓣图书TOP250	京东商品	天气预警				
城市:北京	城市:北京	城市:北京	城市:北京	城市:北京	城市:北京	城市:北京
日期:26日 (今天)	日期:27日 (明天)	日期:28日 (后天)	日期:29日 (周一)	日期:30日 (周二)	日期:1日 (周三)	日期:2日 (周四)
天气:多云	天气:小雨转多云	天气:多云	天气:多云转小雨	天气:多云	天气:多云	天气:晴
气温:10℃	气温:16℃/6℃	气温:20℃/9℃	气温:24℃/13℃	气温:26℃/14℃	气温:25℃/14℃	气温:28℃/15℃
城市:上海	城市:上海	城市:上海	城市:上海	城市:上海	城市:上海	城市:上海
日期:26日 (今天)	日期:27日 (明天)	日期:28日 (后天)	日期:29日 (周一)	日期:30日 (周二)	日期:1日 (周三)	日期:2日 (周四)
天气:晴	天气:多云转小雨	天气:小雨	天气:中雨	天气:阴转多云	天气:多云转阴	天气:阴转多云
气温:11℃	气温:18℃/14℃	气温:22℃/17℃	气温:22℃/17℃	气温:20℃/17℃	气温:25℃/17℃	气温:23℃/11℃
城市:广州	城市:广州	城市:广州	城市:广州	城市:广州	城市:广州	城市:广州
日期:26日 (今天)	日期:27日 (明天)	日期:28日 (后天)	日期:29日 (周一)	日期:30日 (周二)	日期:1日 (周三)	日期:2日 (周四)
天气:中雨	天气:中雨转雷阵雨	天气:雷阵雨	天气:雷阵雨转中雨	天气:中雨转大雨	天气:大雨转多云	天气:多云
气温:24℃	气温:28℃/24℃	气温:29℃/24℃	气温:30℃/25℃	气温:29℃/23℃	气温:26℃/22℃	气温:28℃/21℃
城市:温州	城市:温州	城市:温州	城市:温州	城市:温州	城市:温州	城市:温州
日期:26日 (今天)	日期:27日 (明天)	日期:28日 (后天)	日期:29日 (周一)	日期:30日 (周二)	日期:1日 (周三)	日期:2日 (周四)
天气:阴	天气:阴转小雨	天气:小雨转多云	天气:小雨	天气:小雨	天气:多云	天气:晴