温州大学瓯江学院数信分院

爬虫与数据分析 实验报告

实验名称：	实验二：JSP+css+div 开发首页（一）				
班　　级：	16 计算机四班	姓　　名：	姚周晖	学　　号：	16219111407
实验地点：	B7-403	日　　期：	2019.4.18		

一、实验目的：

1、静态网页爬取
2、动态网页爬取

二、实验环境：

Python，django

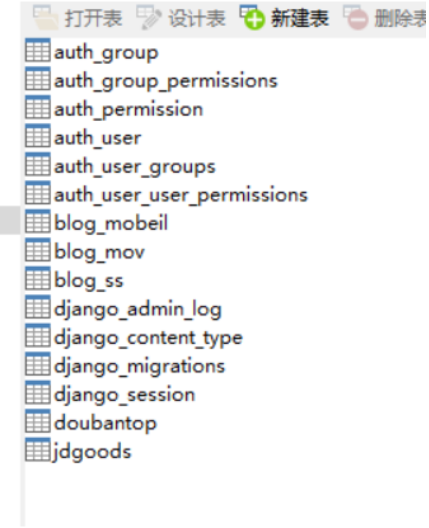
三、爬虫代码：

四.爬取 top250 的电影

```
from bs4 import BeautifulSoup

import pymysql

import requests

import re

import os


def connect_db():

    connect = pymysql.connect(

        user="root",

        password="admin",

        host="localhost",

        db="ojmovie",

        port=3306,

        charset=("utf8"),

        use_unicode=True,

    )

    return connect


def get_html(web_url):

    header = {

        "User-Agent": "Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/47.0.2526.108
Safari/537.36 2345Explorer/8.5.1.15355"}

    html = requests.get(url=web_url, headers=header).text

    Soup = BeautifulSoup(html, "lxml")

    data = Soup.find("ol").find_all("li")
```

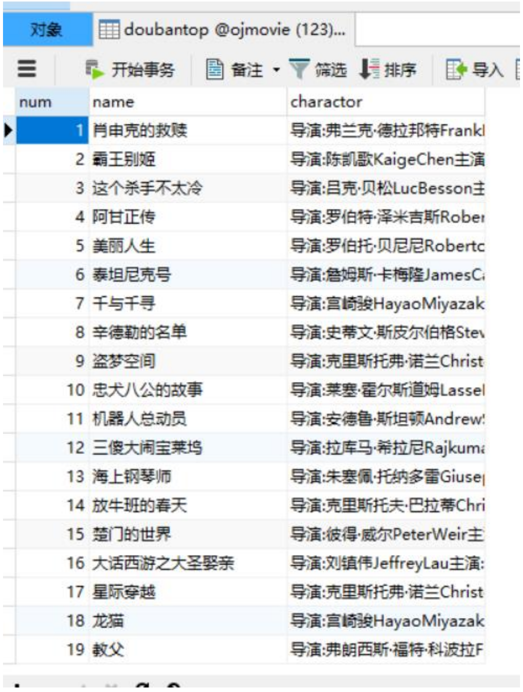
```python
    return data


def get_info(all_move, connect, cursor):
    for info in all_move:

        nums = re.findall(r'<em class="">\d+</em>', str(info), re.S | re.M)
        nums = re.findall(r'\d+', str(nums), re.S | re.M)
        num = nums[0]


        names = info.find("span")
        name = names.get_text()


        charactors = info.find("p")
        charactor = charactors.get_text().replace(" ", "").replace("\n", "")
        charactor = charactor.replace("\xa0", "").replace("\xee", "").replace("\xf6", "").replace("\u0161", "").replace("\xf4",
"").replace("\xfb", "").replace("\u2027", "")


        data = {'num':num, 'name':name, 'charactor':charactor}
        print(data)
        # 保存数据
        cursor.execute("insert into doubantop(num,name,charactor)values(%s,%s,%s)",
                       [data['num'], data['name'], data['charactor']])
        # 提交
        connect.commit()
    return


if __name__ == "__main__":
    connect = connect_db()
    cursor = connect.cursor()
    page = 0
    while page <= 225:
        web_url = "https://movie.douban.com/top250?start=%s&filter=" % page
        all_move = get_html(web_url)
        data = get_info(all_move, connect, cursor)
        page += 25

    connect.close()
```

数据库文件：



auth_group
auth_group_permissions
auth_permission
auth_user
auth_user_groups
auth_user_user_permissions
blog_mobeil
blog_mov
blog_ss
django_admin_log
django_content_type
django_migrations
django_session
doubantop
jdgoods

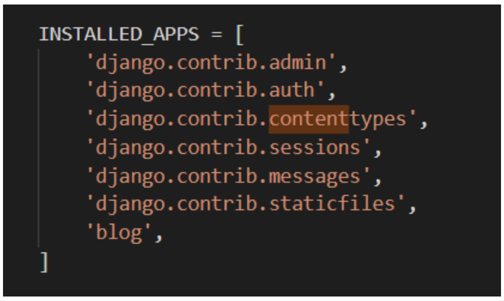
Doubantop：



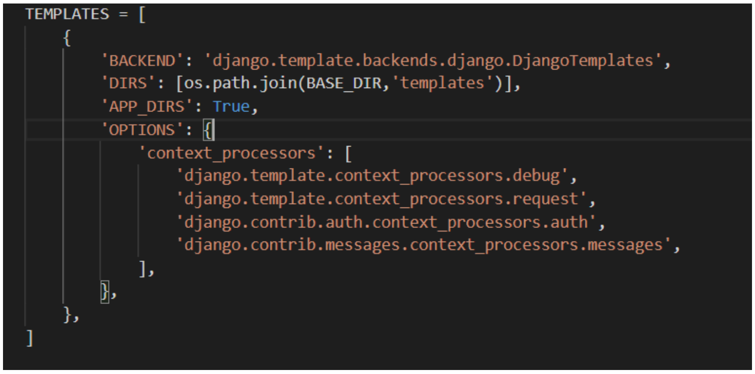
num	name	charactor
1	肖申克的救赎	导演:弗兰克·德拉邦特Frankl
2	霸王别姬	导演:陈凯歌KaigeChen主演
3	这个杀手不太冷	导演:吕克·贝松LucBesson主
4	阿甘正传	导演:罗伯特·泽米吉斯Rober
5	美丽人生	导演:罗伯托·贝尼尼Robertc
6	泰坦尼克号	导演:詹姆斯·卡梅隆JamesC
7	千与千寻	导演:宫崎骏HayaoMiyazak
8	辛德勒的名单	导演:史蒂文·斯皮尔伯格Stev
9	盗梦空间	导演:克里斯托弗·诺兰Christ
10	忠犬八公的故事	导演:莱塞·霍尔斯道姆Lassel
11	机器人总动员	导演:安德鲁·斯坦顿Andrew
12	三傻大闹宝莱坞	导演:拉库马·希拉尼Rajkum
13	海上钢琴师	导演:朱塞佩·托纳多雷Giuse
14	放牛班的春天	导演:克里斯托夫·巴拉蒂Chri
15	楚门的世界	导演:彼得·威尔PeterWeir主
16	大话西游之大圣娶亲	导演:刘镇伟JeffreyLau主演:
17	星际穿越	导演:克里斯托弗·诺兰Christ
18	龙猫	导演:宫崎骏HayaoMiyazak
19	教父	导演:弗朗西斯·福特·科波拉F

使用 django 显示数据：

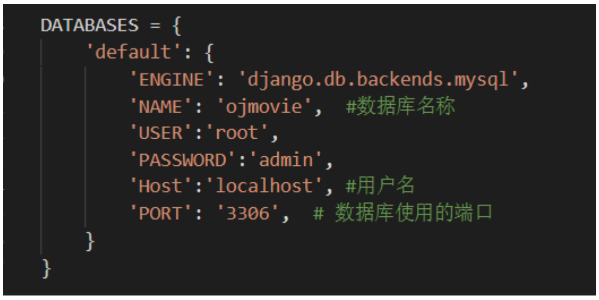
Setting:



```
INSTALLED_APPS = [
    'django.contrib.admin',
    'django.contrib.auth',
    'django.contrib.contenttypes',
    'django.contrib.sessions',
    'django.contrib.messages',
    'django.contrib.staticfiles',
    'blog',
]
```

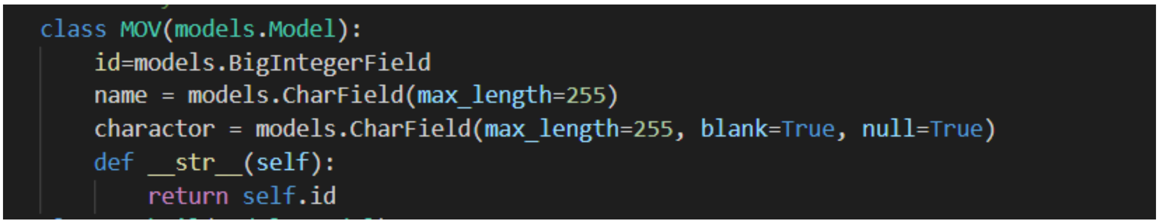


```
TEMPLATES = [
    {
        'BACKEND': 'django.template.backends.django.DjangoTemplates',
        'DIRS': [os.path.join(BASE_DIR,'templates')],
        'APP_DIRS': True,
        'OPTIONS': {
            'context_processors': [
                'django.template.context_processors.debug',
                'django.template.context_processors.request',
                'django.contrib.auth.context_processors.auth',
                'django.contrib.messages.context_processors.messages',
            ],
        },
    },
]
```



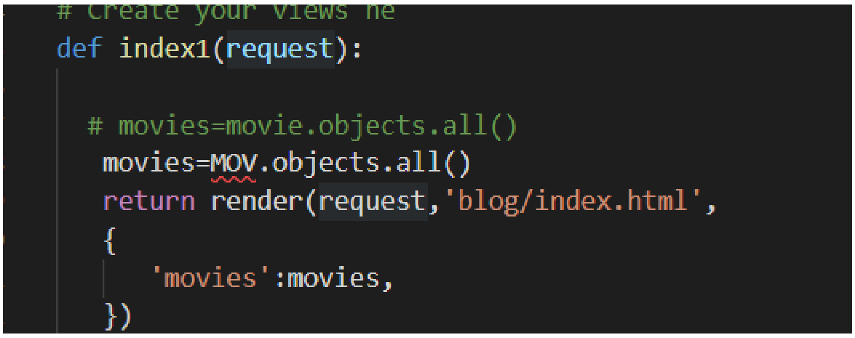
```
DATABASES = {
    'default': {
        'ENGINE': 'django.db.backends.mysql',
        'NAME': 'ojmovie',  #数据库名称
        'USER':'root',
        'PASSWORD':'admin',
        'Host':'localhost', #用户名
        'PORT': '3306',  # 数据库使用的端口
    }
}
```

Models.py:

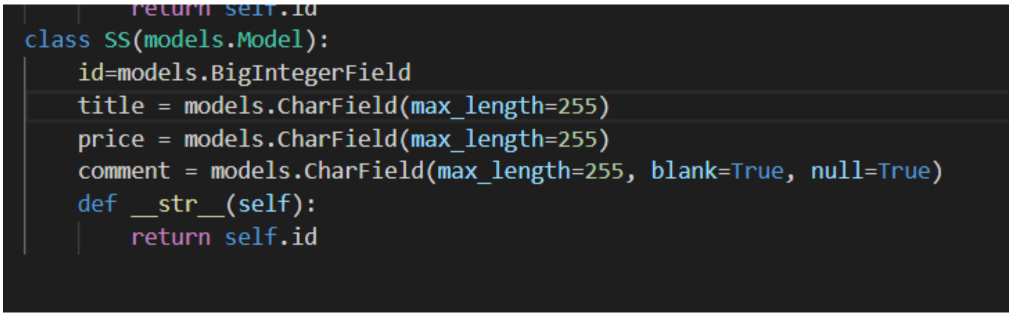


```
class MOV(models.Model):
    id=models.BigIntegerField
    name = models.CharField(max_length=255)
    charactor = models.CharField(max_length=255, blank=True, null=True)
    def __str__(self):
        return self.id
```

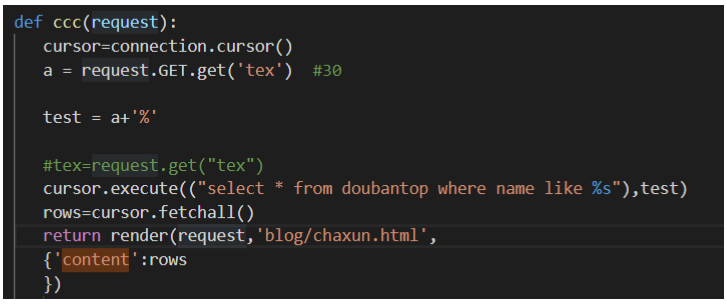
Views.py:



```
# Create your views he
def index1(request):

    # movies=movie.objects.all()
     movies=MOV.objects.all()
     return render(request,'blog/index.html',
     {
         'movies':movies,
     })
```

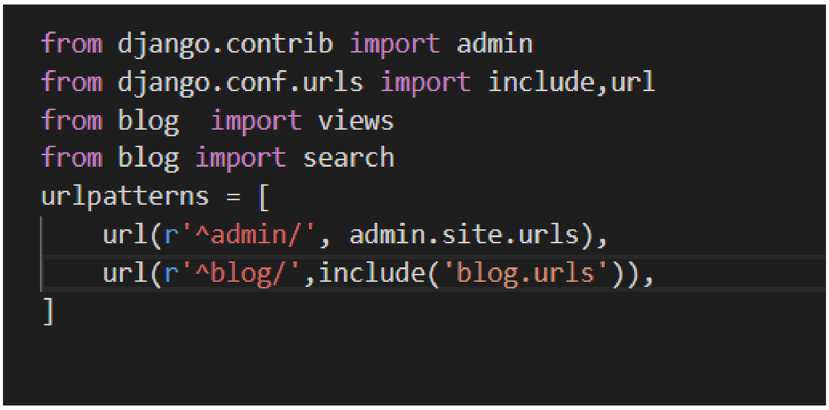


```
        return self.id
class SS(models.Model):
    id=models.BigIntegerField
    title = models.CharField(max_length=255)
    price = models.CharField(max_length=255)
    comment = models.CharField(max_length=255, blank=True, null=True)
    def __str__(self):
        return self.id
```



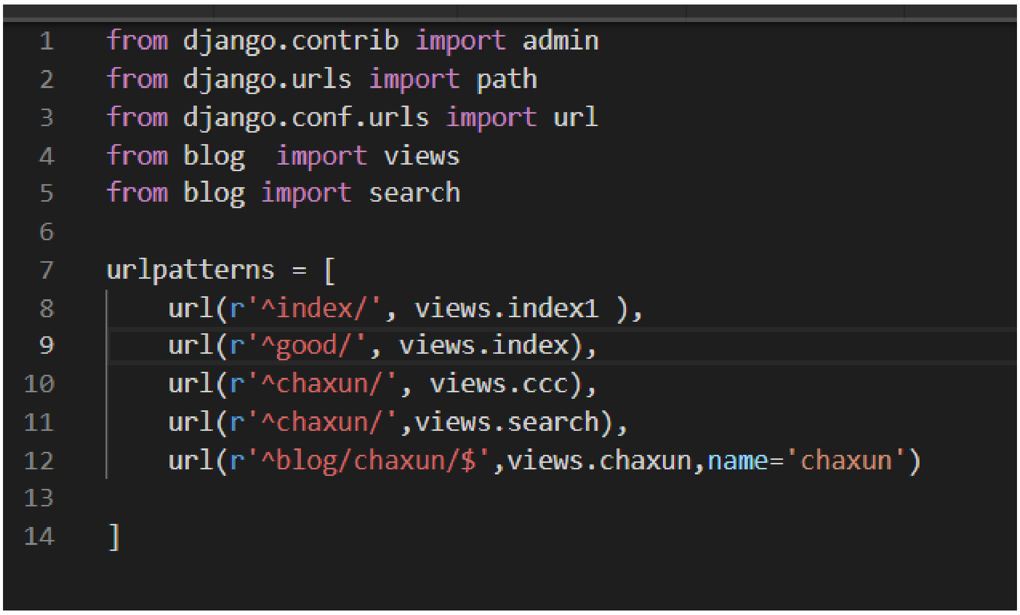
```
def ccc(request):
    cursor=connection.cursor()
    a = request.GET.get('tex')  #30

    test = a+'%'

    #tex=request.get("tex")
    cursor.execute(("select * from doubantop where name like %s"),test)
    rows=cursor.fetchall()
    return render(request,'blog/chaxun.html',
    {'content':rows
    })
```

Urls:



```
from django.contrib import admin
from django.conf.urls import include,url
from blog  import views
from blog import search
urlpatterns = [
    url(r'^admin/', admin.site.urls),
    url(r'^blog/',include('blog.urls')),
]
```

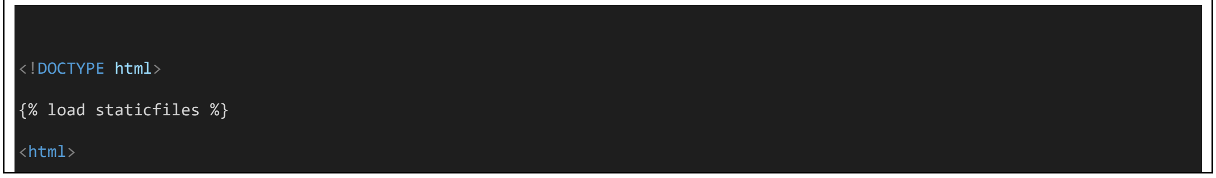
看网上视频将 urls 分写到 blog 项目中：（blog.urls.py）



```
from django.contrib import admin
from django.urls import path
from django.conf.urls import url
from blog  import views
from blog import search

urlpatterns = [
    url(r'^index/', views.index1 ),
    url(r'^good/', views.index),
    url(r'^chaxun/', views.ccc),
    url(r'^chaxun/',views.search),
    url(r'^blog/chaxun/$',views.chaxun,name='chaxun')

]
```

页面设计（主页和搜索页）：

Index.html:



```
<!DOCTYPE html>

{% load staticfiles %}

<html>
```

```
<head>
    <title>MyBlog</title>
    <link rel="stylesheet" href="{% static 'css/main.css' %}"/>
    <style>
            body{
               background-color: yellow;
            }
            b{

                float: right;;

            }
    </style>
</head>

<body >
    <div id="main"><img src="{% static "img/5.jpg" %}" width="100%" height="200px"></div>
    <form action="http://127.0.0.1:8000/blog/chaxun/" method="GET"> <!-- 将 http 这修改为后台地址  method get/post 与后台的要一致
-->
    <input type="text" name='tex'> <!-- 这里就收参数 name 要与后台一致 -->

    <input type="submit" value="搜索">
    </form>
            <br>
            <br>
        <center>
            <table border="1px">

                <tr>

                    <th>排名</th>
                    <th>电影名</th>
                    <th>导演</th>
                    <th>前往观看</th>



                </tr>

            {% for new  in  movies %}
                <tr>
                    <td><center>{{new.id}}</center></td>
                    <td><center>{{new.name}}</center></td>
                    <td><center>{{new.charactor}}</center></td>
                    <th><center><a href="https://movie.douban.com/">点击观看</a></center></th>
                </tr>
```

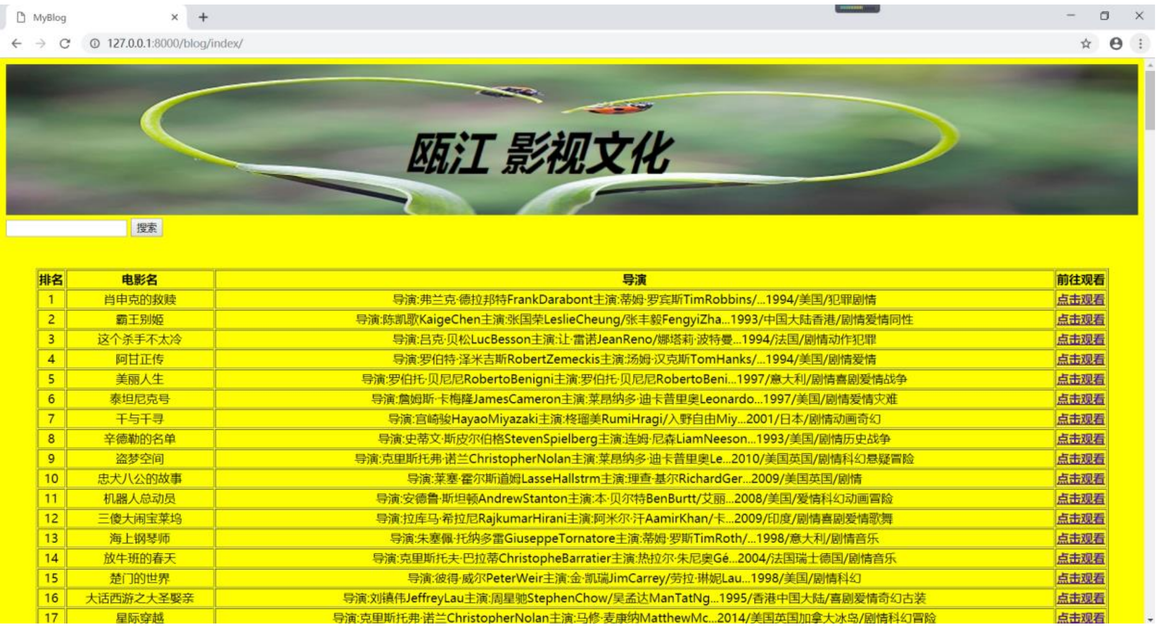
```
           {% endfor%}
             </table>
           </center>



</body>

</html>
```

Chaxun.html:

```
<!DOCTYPE html>
{% load staticfiles %}
<html>

<head>
    <title>MyBlog</title>
    <link rel="stylesheet" href="{% static 'css/main.css' %}"/>
    <style>
           body{
              background-color: yellow;
           }
           label{

               float: right;

           }
    </style>
</head>

<body >

    <div id="main"><img src="{% static "img/5.jpg" %}" width="100%" height="200px"></div>
    <br>
    <br>
    <laebl>
           <a href="http://127.0.0.1:8000/blog/index/">
               <input type="button" value="返回首页">
           </a>
    </label>
    <br>
        <center> <font size="50px">   <tr>搜索结果</tr></font></center>
           <br>
           <br>
```

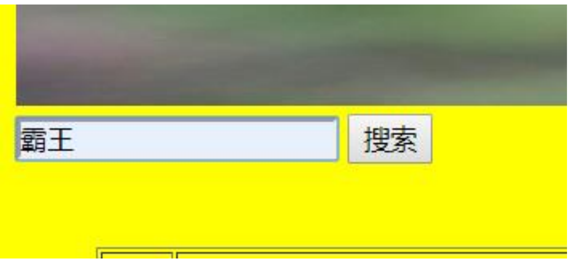
```
                <tr>

                        <center> {{content}} <a href="https://movie.douban.com/">点击观看</a></center>

                </tr>




</body>

</html>
```

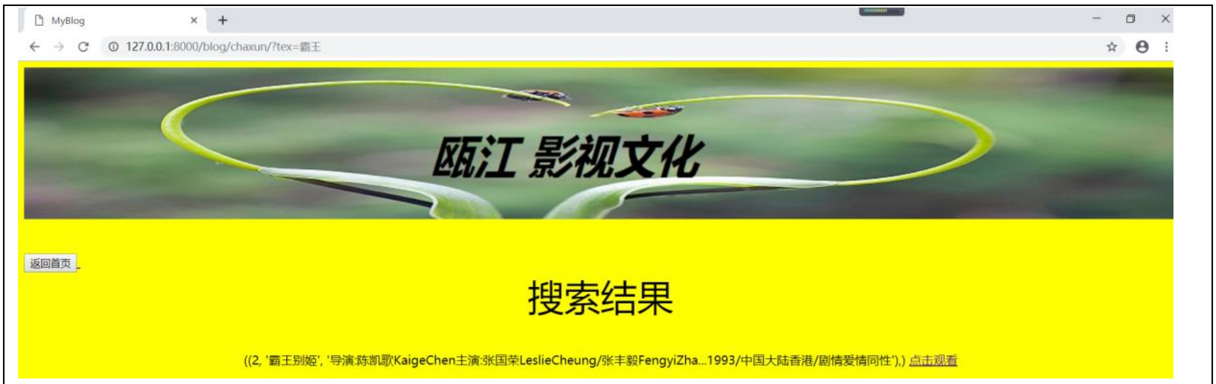
效果展示：



排名	电影名	导演	前往观看
1	肖申克的救赎	导演:弗兰克·德拉邦特FrankDarabont主演:蒂姆·罗宾斯TimRobbins/...1994/美国/犯罪剧情	点击观看
2	霸王别姬	导演:陈凯歌KaigeChen主演:张国荣LeslieCheung/张丰毅FengyiZha...1993/中国大陆香港/剧情爱情同性	点击观看
3	这个杀手不太冷	导演:吕克·贝松LucBesson主演:让·雷诺JeanReno/娜塔莉·波特曼...1994/法国/剧情动作犯罪	点击观看
4	阿甘正传	导演:罗伯特·泽米吉斯RobertZemeckis主演:汤姆·汉克斯TomHanks/...1994/美国/剧情爱情	点击观看
5	美丽人生	导演:罗伯托·贝尼尼RobertoBenigni主演:罗伯托·贝尼尼RobertoBeni...1997/意大利/剧情喜剧爱情战争	点击观看
6	泰坦尼克号	导演:詹姆斯·卡梅隆JamesCameron主演:莱昂纳多·迪卡普里奥Leonardo...1997/美国/剧情爱情灾难	点击观看
7	千与千寻	导演:宫崎骏HayaoMiyazaki主演:柊瑠美RumiHragi/入野自由Miy...2001/日本/剧情动画奇幻	点击观看
8	辛德勒的名单	导演:史蒂文·斯皮尔伯格StevenSpielberg主演:连姆·尼森LiamNeeson...1993/美国/剧情历史战争	点击观看
9	盗梦空间	导演:克里斯托弗·诺兰ChristopherNolan主演:莱昂纳多·迪卡普里奥Le...2010/美国英国/剧情科幻悬疑冒险	点击观看
10	忠犬八公的故事	导演:莱塞·霍尔斯道姆LasseHallstrm主演:理查·基尔RichardGer...2009/美国英国/剧情	点击观看
11	机器人总动员	导演:安德鲁·斯坦顿AndrewStanton主演:本·贝尔特BenBurtt/艾丽...2008/美国/爱情科幻动画冒险	点击观看
12	三傻大闹宝莱坞	导演:拉库马·希拉尼RajkumarHirani主演:阿米尔·汗AamirKhan/卡...2009/印度/剧情喜剧爱情歌舞	点击观看
13	海上钢琴师	导演:朱塞佩·托纳多雷GiuseppeTornatore主演:蒂姆·罗斯TimRoth/...1998/意大利/剧情音乐	点击观看
14	放牛班的春天	导演:克里斯托夫·巴拉蒂ChristopheBarratier主演:热拉尔·朱尼奥Gé...2004/法国瑞士德国/剧情音乐	点击观看
15	楚门的世界	导演:彼得·威尔PeterWeir主演:金·凯瑞JimCarrey/劳拉·琳妮Lau...1998/美国/剧情科幻	点击观看
16	大话西游之大圣娶亲	导演:刘镇伟JeffreyLau主演:周星驰StephenChow/吴孟达ManTatNg...1995/香港中国大陆/喜剧爱情奇幻古装	点击观看
17	星际穿越	导演:克里斯托弗·诺兰ChristopherNolan主演:马修·麦康纳MatthewMc...2014/美国英国加拿大冰岛/剧情科幻冒险	点击观看

搜索功能：





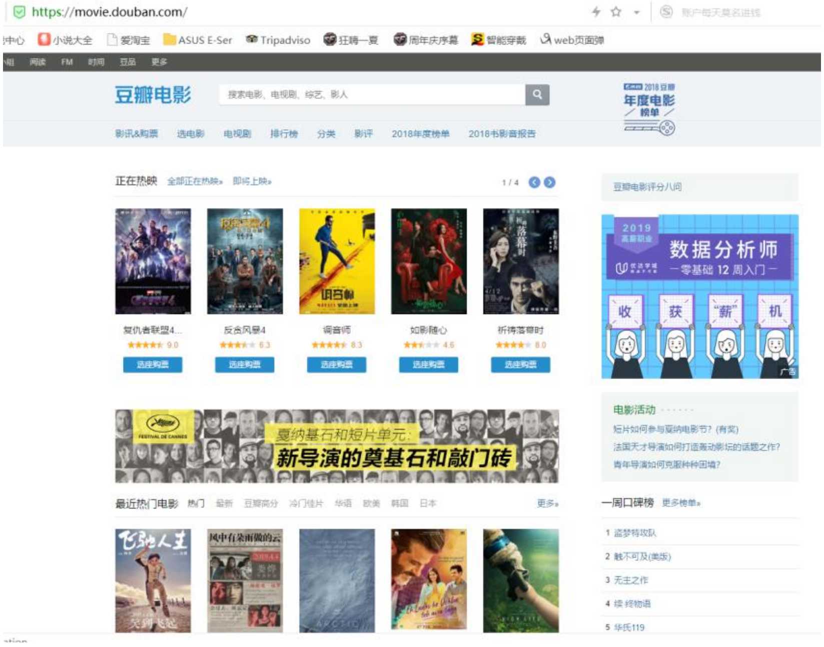
搜索结果

((2, '霸王别姬', '导演:陈凯歌KaigeChen主演:张国荣LeslieCheung/张丰毅FengyiZha...1993/中国大陆香港/剧情爱情同性'),) 点击观看

实现的功能：

爬取了 top250 并写入了数据库，然后在 django 中显示。设计了 css，提高页面友好度，点击观看跳转至电影页面；增加搜索功能，根据电影的部分关键名。

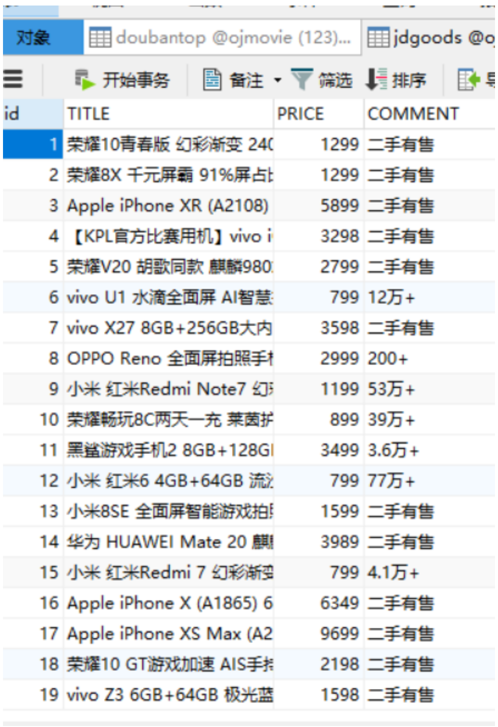
如：点击后跳转：



五、动态爬取京东手机产品(装了谷歌的驱动)：

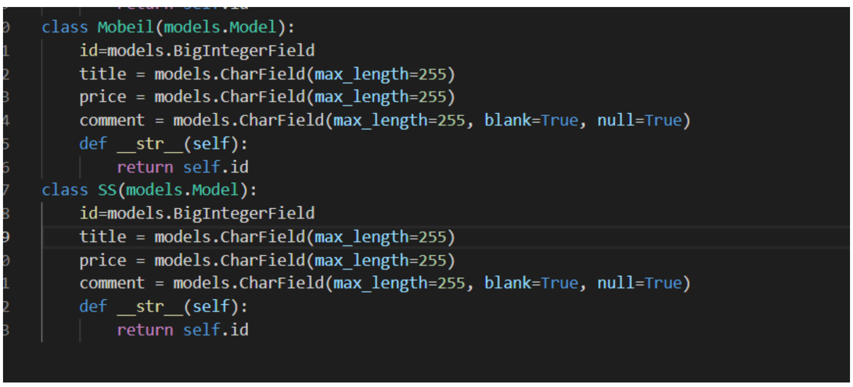
```
from selenium.webdriver import Chrome

from config import *

from selenium.webdriver.support.ui import WebDriverWait

from selenium.webdriver.support import expected_conditions as EC

from selenium.webdriver.common.by import By

from selenium.webdriver.common.keys import Keys

from time import sleep

from selenium.common.exceptions import NoSuchElementException

import pymysql


def next_page(client, wait, page_num):
    END_PAGE = 1
    while len(client.find_elements_by_class_name('gl-item')) < 60:
        client.execute_script('window.scrollTo(0, document.body.scrollHeight)')
        sleep(1)
    print("[+] 第{}加载完成".format(page_num))

    parse_page(page_num, client)

    # 下一页
    page_num += 1
    if page_num > END_PAGE:
        print('前{}页爬取成功'.format(END_PAGE))
        return


    wait.until(
        EC.presence_of_element_located(
            (By.CSS_SELECTOR, '#J_bottomPage > span.p-skip > input')
        )
    )

    wait.until(
        EC.element_to_be_clickable(
            (By.CSS_SELECTOR, '#J_bottomPage > span.p-skip > a')
        )
    )

    input_ = client.find_element_by_css_selector('#J_bottomPage > span.p-skip > input')
    input_.clear()
    input_.send_keys(page_num)

    input_.send_keys(Keys.ENTER)
    wait.until(
        EC.text_to_be_present_in_element(
            (By.CSS_SELECTOR, '#J_bottomPage > span.p-num > a.curr'),
            str(page_num)
        )
    )

    next_page(client, wait, page_num)

def connect_db():
    connect = pymysql.connect(
        user="root",
        password="admin",
        host="localhost",
        db="ojmovie",
        port=3306,
        charset=("utf8"),
        use_unicode=True,
    )
    return connect
def parse_page(page_num, client):
    print("[+] 开始解析第{}页数据".format(page_num))
    items = client.find_elements_by_class_name('gl-item')
    index = 1

    for item in items:
        print("[{}] ".format(index), end="")
        try:
            title = item.find_element_by_css_selector("div.p-name > a > em").text
        except NoSuchElementException:
            title = None
        try:
            price = item.find_element_by_css_selector("div.p-price > strong > i").text
        except NoSuchElementException:
            price = None


        try:
            comment = item.find_element_by_css_selector(".p-commit a").text
        except NoSuchElementException:
            comment = None

        print("{} >>> {}   >>> {}".format(title, price, comment))

        connect=connect_db()
        cursor = connect.cursor()
        cursor.execute("insert into jdgoods(id,TITLE,PRICE,COMMENT)values (%s,%s,%s,%s)",(index,title, price, comment))
        connect.commit()
        index += 1
    print("[+] 解析第{}页数据完成".format(page_num))
    connect.close()

def search(client, url, keyword,wait):
    client.get(url)
    wait.until(
        EC.presence_of_element_located(
            (By.ID, 'key')
        )
    )
    wait.until(
        EC.element_to_be_clickable(
            (By.CSS_SELECTOR, '#search > div > div.form > button > i')
        )
    )



    input_ = client.find_element_by_id('key')
    input_.send_keys(keyword)



    botton = client.find_element_by_css_selector('#search > div > div.form > button > i')
    botton.click()
    print("[+] 点击搜索完成")

    # 翻页
    page_num = 1
    next_page(client, wait, page_num)

def main():


    client = Chrome()
    url = "http://www.jd.com"

    KEYWORD = '手机'

    wait = WebDriverWait(client, 10)
    search(client, url, KEYWORD, wait)

if __name__ == '__main__':
    main()
```

Jdgoods:



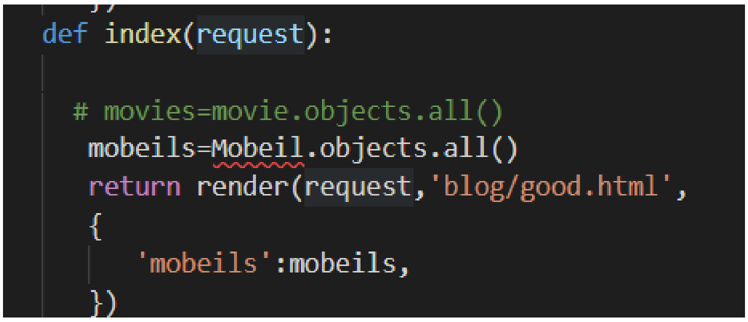
id	TITLE	PRICE	COMMENT
1	荣耀10青春版 幻彩渐变 24	1299	二手有售
2	荣耀8X 千元屏霸 91%屏占	1299	二手有售
3	Apple iPhone XR (A2108)	5899	二手有售
4	【KPL官方比赛用机】vivo i	3298	二手有售
5	荣耀V20 胡歌同款 麒麟980	2799	二手有售
6	vivo U1 水滴全面屏 AI智慧	799	12万+
7	vivo X27 8GB+256GB大内	3598	二手有售
8	OPPO Reno 全面屏拍照手	2999	200+
9	小米 红米Redmi Note7 幻	1199	53万+
10	荣耀畅玩8C两天一充 莱茵	899	39万+
11	黑鲨游戏手机2 8GB+128G	3499	3.6万+
12	小米 红米6 4GB+64GB 流	799	77万+
13	小米8SE 全面屏智能游戏拍	1599	二手有售
14	华为 HUAWEI Mate 20 麒	3989	二手有售
15	小米 红米Redmi 7 幻彩渐	799	4.1万+
16	Apple iPhone X (A1865) 6	6349	二手有售
17	Apple iPhone XS Max (A2	9699	二手有售
18	荣耀10 GT游戏加速 AIS手	2198	二手有售
19	vivo Z3 6GB+64GB 极光蓝	1598	二手有售

Models.py:



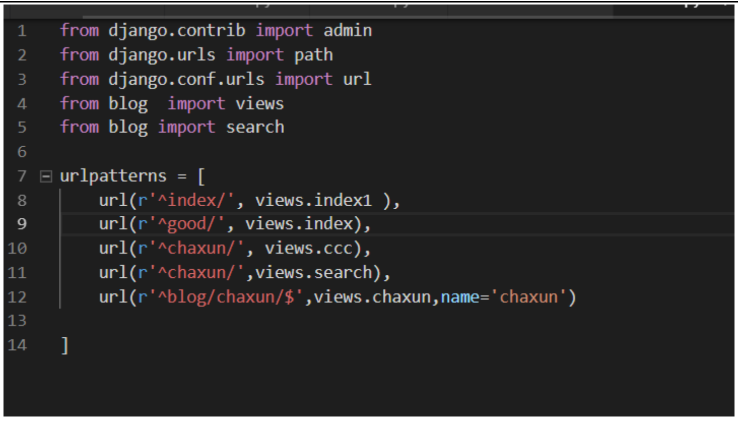
```
class Mobeil(models.Model):
    id=models.BigIntegerField
    title = models.CharField(max_length=255)
    price = models.CharField(max_length=255)
    comment = models.CharField(max_length=255, blank=True, null=True)
    def __str__(self):
        return self.id
class SS(models.Model):
    id=models.BigIntegerField
    title = models.CharField(max_length=255)
    price = models.CharField(max_length=255)
    comment = models.CharField(max_length=255, blank=True, null=True)
    def __str__(self):
        return self.id
```

Views,py:



```
def index(request):

   # movies=movie.objects.all()
    mobeils=Mobeil.objects.all()
    return render(request,'blog/good.html',
    {
        'mobeils':mobeils,
    })
```

Urls.py:

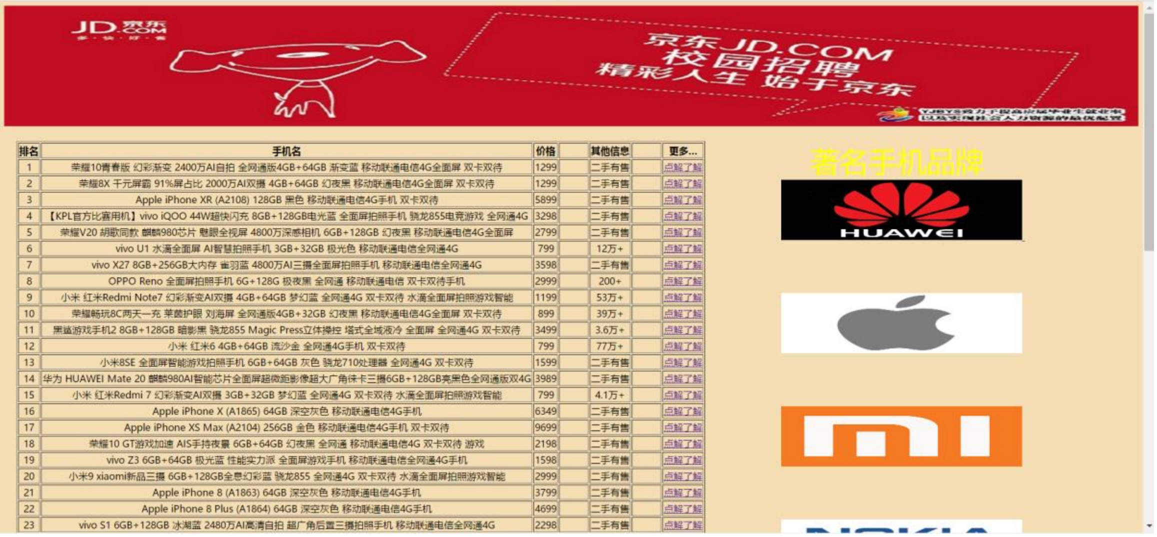


```
from django.contrib import admin
from django.urls import path
from django.conf.urls import url
from blog  import views
from blog import search

urlpatterns = [
    url(r'^index/', views.index1 ),
    url(r'^good/', views.index),
    url(r'^chaxun/', views.ccc),
    url(r'^chaxun/',views.search),
    url(r'^blog/chaxun/$',views.chaxun,name='chaxun')

]
```

页面设计（good.html）：

```
<!DOCTYPE html>

{% load staticfiles %}

<html>


<head>

    <title>MyBlog</title>

    <style>

            body{

                background-color: wheat;






            }

            table{

               margin-left: 20px;

                float: left;






            }

            label

            {   color: yellow;

                float:right;

                margin-right: 10%;




            }



    </style>

</head>



<body>

        <div id="a">

            <img src="{% static "img/4.jpg" %}" width="100%" height="200px">

            <br/>

            <br/>
        </div>


           <table  border="1px" >

                   <tr>

                      <th>排名</th>
                      <th>手机名</th>
                      <th>价格</th>
                      <th>&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;</th>
                      <th>其他信息</th>
                      <th>&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;</th>
                      <th>更多...</th>

                      </tr>

                      {% for a  in  mobeils %}
                          <tr>
                              <td><center>{{a.id}}</center></td>
                              <td><center>{{a.title}}</center></td>
                              <td><center>{{a.price}}</center></td>
                              <td>&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;</td>
                              <td><center>{{a.comment}}</center></td>
                              <td>&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;</td>
                              <td><a
href="https://search.jd.com/Search?keyword=%E6%89%8B%E6%9C%BA&enc=utf-8&wq=%E6%89%8B%E6%9C%BA&pvid=d8c7439106bc4b40939fc102b20246df">点解了解</td>

                          </tr>
                      {% endfor%}

                      </table>
            <label>
               &nbsp;&nbsp;&nbsp;  &nbsp;&nbsp;&nbsp;<font size="100px" color:yellow> 著名手机品牌</font>
              <br/>
              <a href="https://www.vmall.com/?cid=91895">
              <img src="{% static "img/huawei.jpeg" %}"  width="400px" height="100px">
              </a>
              <br/><br/><br/><br/><br/>
              <a href="https://www.apple.com/cn/?afid=p238%7C11DPVC7K_mtid_18707vxu38484&cid=aos-cn-kwsg-brand"></a>
              <img src="{% static "img/iphone.jpg" %}" width="400px" height="100px">
              </a>
              <br/><br/><br/><br/><br/>
              <a href="https://www.mi.com/"></a>
              <img src="{% static "img/xiaomi.png" %}" width="400px" height="100px">

              </a>

              <br/><br/><br/><br/><br/>

              <a href="https://www.nokia.com/zh_int/"></a>

              <img src="{% static "img/rjy.jpg" %}" width="400px" height="100px">

              </a>

            </label>



  </body>



  </html>
```

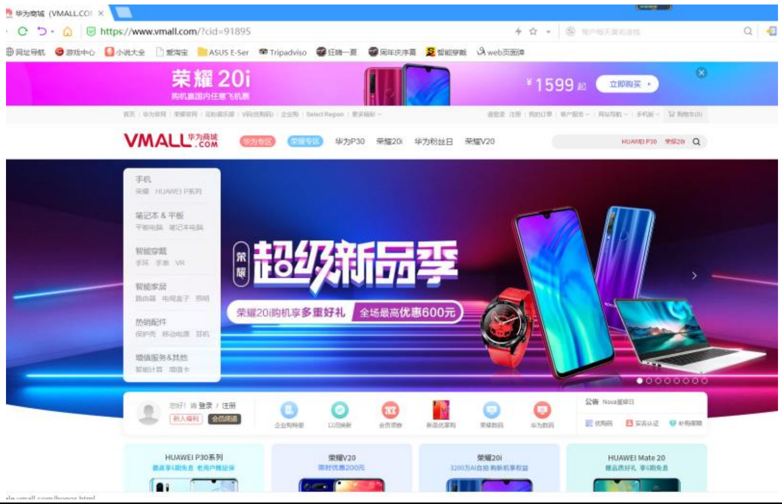
效果展示：



排名	手机名	价格		其他信息		更多...
1	荣耀10青春版 幻彩渐变 2400万AI自拍 全网通版4GB+64GB 渐变蓝 移动联通电信4G全面屏 双卡双待	1299		二手有售		点解了解
2	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	1299		二手有售		点解了解
3	Apple iPhone XR (A2108) 128GB 黑色 移动联通电信4G手机 双卡双待	5899		二手有售		点解了解
4	【KPL官方比赛用机】vivo iQOO 44W超快闪充 8GB+128GB电光蓝 全面屏拍照手机 骁龙855电竞游戏 全网通4G	3298		二手有售		点解了解
5	荣耀V20 胡歌同款 麒麟980芯片 魅眼全视屏 4800万深感相机 6GB+128GB 幻夜黑 移动联通电信4G全面屏	2799		二手有售		点解了解
6	vivo U1 水滴全面屏 AI智慧拍照手机 3GB+32GB 极光色 移动联通电信全网通4G	799		12万+		点解了解
7	vivo X27 8GB+256GB大内存 雀羽蓝 4800万AI三摄全面屏拍照手机 移动联通电信全网通4G	3598		二手有售		点解了解
8	OPPO Reno 全面屏拍照手机 6G+128G 极夜黑 全网通 移动联通电信 双卡双待手机	2999		200+		点解了解
9	小米 红米Redmi Note7 幻彩渐变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	1199		53万+		点解了解
10	荣耀畅玩8C两天一充 莱茵护眼 刘海屏 全网通版4GB+32GB 幻夜黑 移动联通电信4G全面屏 双卡双待	899		39万+		点解了解
11	黑鲨游戏手机2 8GB+128GB 暗影黑 骁龙855 Magic Press立体操控 塔式全域液冷 全面屏 全网通4G 双卡双待	3499		3.6万+		点解了解
12	小米 红米6 4GB+64GB 流沙金 全网通4G手机 双卡双待	799		77万+		点解了解
13	小米8SE 全面屏智能游戏拍照手机 6GB+64GB 灰色 骁龙710处理器 全网通4G 双卡双待	1599		二手有售		点解了解
14	华为 HUAWEI Mate 20 麒麟980AI智能芯片全面屏超微距影像超大广角徕卡三摄6GB+128GB亮黑色全网通版双4G	3989		二手有售		点解了解
15	小米 红米Redmi 7 幻彩渐变AI双摄 3GB+32GB 梦幻蓝 全网通4G 双卡双待 水滴全面屏拍照游戏智能	799		4.1万+		点解了解
16	Apple iPhone X (A1865) 64GB 深空灰色 移动联通电信4G手机	6349		二手有售		点解了解
17	Apple iPhone XS Max (A2104) 256GB 金色 移动联通电信4G手机 双卡双待	9699		二手有售		点解了解
18	荣耀10 GT游戏加速 AIS手持夜景 6GB+64GB 幻夜黑 全网通 移动联通电信4G 双卡双待 游戏	2198		二手有售		点解了解
19	vivo Z3 6GB+64GB 极光蓝 性能实力派 全面屏游戏手机 移动联通电信全网通4G手机	1598		二手有售		点解了解
20	小米9 xiaomi新品三摄 6GB+128GB全息幻彩蓝 骁龙855 全网通4G 双卡双待 水滴全面屏拍照游戏智能	2999		二手有售		点解了解
21	Apple iPhone 8 (A1863) 64GB 深空灰色 移动联通电信4G手机	3799		二手有售		点解了解
22	Apple iPhone 8 Plus (A1864) 64GB 深空灰色 移动联通电信4G手机	4699		二手有售		点解了解
23	vivo S1 6GB+128GB 冰湖蓝 2480万AI高清自拍 超广角后置三摄拍照手机 移动联通电信全网通4G	2298		二手有售		点解了解

实现功能：使用了 selenium 自动爬取了京东商城的手机部分，并且可以根据更改 KEYWORD 和 page_num，来实现爬取的对象变更和爬取的数量变更，使用 css 提高页面友好度。并设置超链接跳转到对应的手机页面。

如点击了华为图标：


https://www.vmall.com/?cid=91895
荣耀 20i
¥1599 起
立即购买
VMALL.COM 华为商城
华为P30
荣耀20i
华为粉丝日
荣耀V20
手机
笔记本 & 平板
智能穿戴
智能家居
热销配件
增值服务&其他
超级新品季
荣耀20i购机享多重好礼
全场最高优惠600元
HUAWEI P30系列
荣耀V20
荣耀20i
HUAWEI Mate 20