

页面

爬虫展示

爬取京东网站



爬取豆瓣网站

```
1 https://movie.douban.com/top250?start=0&filter=
[肖申克的救赎,霸王别姬,这个杀手不太冷,阿甘正传,美丽人生,泰坦尼克号,千与千寻,辛德勒的名单,盗梦空间,忠犬八公的故事,机器人总动员,三傻大闹宝莱坞,海上钢琴师,放牛班的春天,慧门的世界,大话西游之大圣娶亲,星际穿越,龙猫,教父,肖申克,无间道,疯狂动物城,当幸福来敲门,怦然心动,触不可及]
2 https://movie.douban.com/top250?start=25&filter=
[蝙蝠侠:黑暗骑士,乱世佳人,活着,少年派的奇幻漂流,拉方证人,天堂电影院,鬼子来了,指环王3:王者无敌,十二怒汉,天空之城,排班吧!卷福,飞屋环游记,大话西游之月光宝盒,搏击俱乐部,罗马假日,哈尔的移动城堡,南茜追女人,辩护人,窃听风暴,末代皇帝,两杆大烟枪,飞越疯人院,死亡诗社,素媛,寻梦环游记]
3 https://movie.douban.com/top250?start=50&filter=
[指环王2:双塔奇兵,教父2,N字仇杀队,指环王1:魔戒再现,海豚湾,饮食男女,狮子王,美丽心灵,情书,钢琴家,本杰明 巴顿奇事,美国往事,看不见的客人,燕喜帝国,小鞋子,西西里的美丽传说,大闹天宫,哈利 波特与魔法石,让子弹飞,拯救大兵瑞恩,致命魔术,七宗罪,被嫌弃的松子的一生,音乐之声,低俗小说]
4 https://movie.douban.com/top250?start=75&filter=
[沉默的羔羊,天使爱美丽,猫鼠游戏,勇敢的心,蝴蝶效应,剪刀手爱德华,春光乍泄,心灵捕手,禁闭岛,布达佩斯大饭店,穿条纹睡衣的男孩,久负盛名,阿凡达,幽灵公主,阳光灿烂的日子,致命ID,第六感,加勒比海盗,狩猎,断背山,钢琴师和马克思,重庆森林,摩登时代,喜剧之王,告白]
5 https://movie.douban.com/top250?start=100&filter=
```

Settings 数据库连接

```
DATABASES = {
    'default': {
        'ENGINE': 'django.db.backends.mysql',
        'NAME': 'chd' ,
        'USER': 'root',
        'PASSWORD': 'Chenhaodong123@',
        'HOST': '127.0.0.1',
        'PORT': '3306'
    }
}
```

Settings 想改 en-us 改成 ch-hangs 改了就报错就没改

Urls

```
1 from django.conf.urls import url
2
3 from . import views
4
5 urlpatterns = [
6     url(r'^index/$', views.index),
7 ]
```

```
settings.py x admin.py urls.py chd x wsgi.py views.py
9 Class-based views
10     1. Add an import: from other_app.views import Home
11     2. Add a URL to urlpatterns: path('', Home.as_view(), name='home')
12 Including another URLconf
13     1. Import the include() function: from django.urls import include
14     2. Add a URL to urlpatterns: path('blog/', include('blog.urls'))
15 """
16 from django.contrib import admin
17 from django.conf.urls import url, include
18
19
20
21 urlpatterns = [
22     url(r'^admin/', admin.site.urls),
23     url(r'^chd/', include('yuApp.urls')),
24 ]
25
```

Models

```
1 from __future__ import unicode_literals
2
3 from django.db import models
4
5
6 # Create your models here.
7 class art(models.Model):
8     url =models.CharField(max_length=200,default='')
9     content=models.CharField(max_length=1000,default='')
```

0001:

```

1  # Generated by Django 2.2 on 2019-04-27 02:30
2
3  from django.db import migrations, models
4
5
6  class Migration(migrations.Migration):
7
8      initial = True
9
10     dependencies = [
11     ]
12
13     operations = [
14         migrations.CreateModel(
15             name='art',
16             fields=[
17                 ('id', models.AutoField(auto_created=True, primary_key=True,
18                 ('url', models.CharField(default='', max_length=200,
19                 ('content', models.CharField(default='', max_length=200,
20             ],
21         ),
22     ],
23

```

Views

```

1  from django.shortcuts import render
2  from django.http import HttpResponse
3  from . import models
4
5
6  def index(request):
7      pys=models.art.objects.all()
8      return render(request,'order_list.html',{'pys':pys})

```

Html

```

<head>
  <meta charset="utf-8">
  <title>Bootstrap 实例 - 一个简单的网页</title>
  <link rel="stylesheet"
href="https://cdn.staticfile.org/twitter-bootstrap/3.3.7/css/bootstrap.
min.css">

```

```
<script
src="https://cdn.staticfile.org/jquery/2.1.1/jquery.min.js"></script>
<script
src="https://cdn.staticfile.org/twitter-bootstrap/3.3.7/js/bootstrap.mi
n.js"></script>
<style>
    .fakeimg {
        height: 200px;
        background: #aaa;
    }
</style>
</head>
<body>
<div class="jumbotron text-center" style="margin-bottom:0">
    <h1>爬虫展示</h1>
</div>

<nav class="navbar navbar-inverse">
    <div class="container-fluid">
        <div class="navbar-header">
            <button type="button" class="navbar-toggle" data-toggle="collapse"
data-target="#myNavbar">
                <span class="icon-bar"></span>
                <span class="icon-bar"></span>
                <span class="icon-bar"></span>
            </button>
            <a class="navbar-brand" href=" " >爬虫</a>
        </div>
        <div class="collapse navbar-collapse" id="myNavbar">
            <ul class="nav navbar-nav">
                <li class="active">主页</a></li>
            </ul>
        </div>
    </div>
</nav>

<div class="container">
    <div class="row">
        <div class="col-sm-4">
            <h2>爬取京东网站</h2>
            <div class="fakeimg">

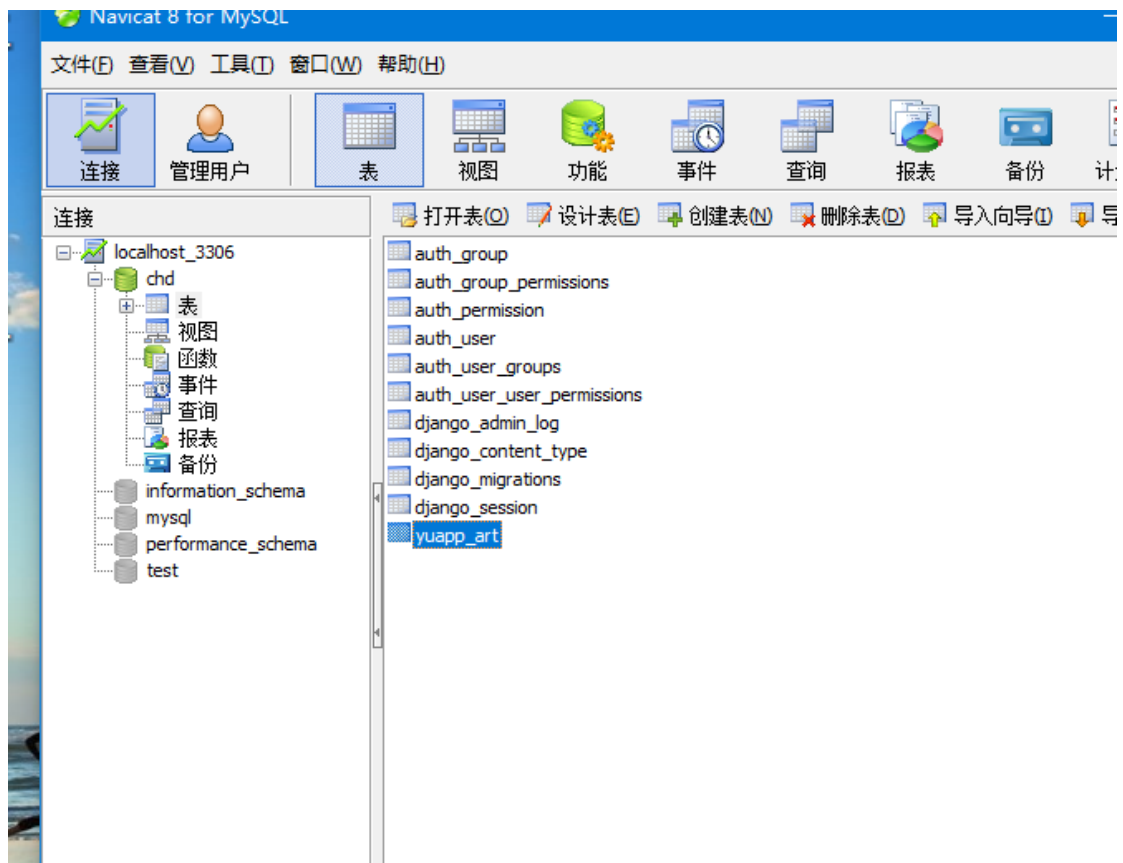
            </div>
            <hr class="hidden-sm hidden-md hidden-lg">
        </div>
    </div>
</div>
```

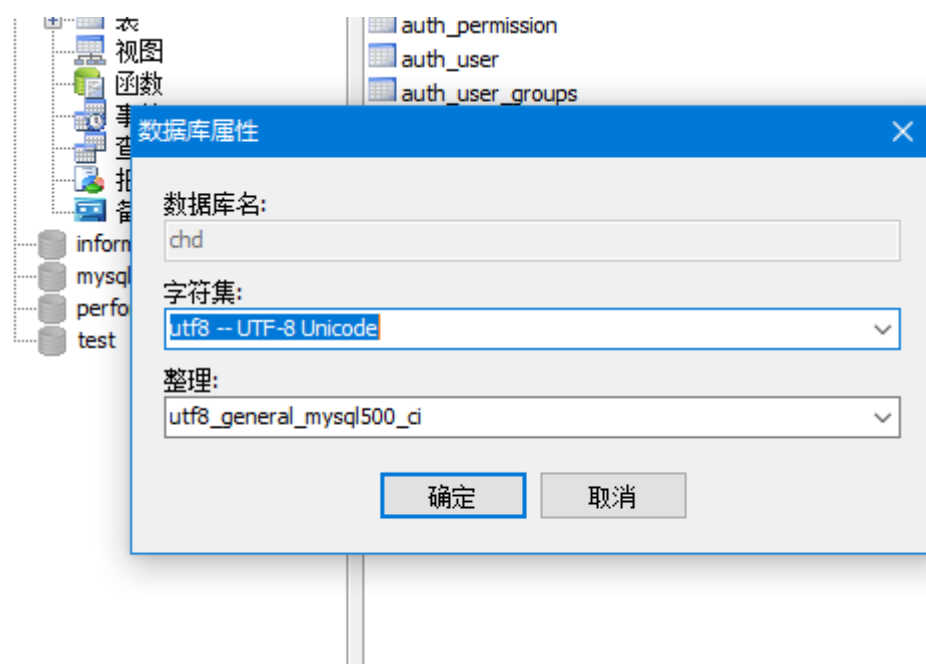
```

        </div>
<div class="col-sm-8">
    <h2>爬取豆瓣网站</h2>
    <div class="fakeimg">
        {% for py in pys %}
        <tr>
            <td>{{py.id}}</td>
            <td>{{py.url}}</td>
            <td>{{py.content}}</td>
        </tr>
        {% endfor %}
    </div>
<br>
<!DOCTYPE html>
<html>

```

数据库





导入向导① 导出向导② 筛检向导 网格视图 表单视图 备注 十六进制 图像		
	url	content
1	https://movie.douban.com/top	[肖申克的救赎, '霸王别姬',
2	https://movie.douban.com/top	[蝙蝠侠：黑暗骑士', '乱世佳
3	https://movie.douban.com/top	[指环王2：双塔奇兵', '教父：
4	https://movie.douban.com/top	[沉默的羔羊', '天使爱美丽',
5	https://movie.douban.com/top	[大鱼', '消失的爱人', '一一',
6	https://movie.douban.com/top	[借东西的小人阿莉埃蒂', '崖
7	https://movie.douban.com/top	[纵横四海', '雨人', '教父3', '云
8	https://movie.douban.com/top	[未麻的部屋', '穿越时空的少
9	https://movie.douban.com/top	[惊魂记', '青蛇', '疯狂的石头
10	https://movie.douban.com/top	[E.T. 外星人', '末路狂花', '聚