



南昌大学实验报告

学生姓名：____丁俊____ 学 号：____8003119100____ 专业班级：____信安 193 班____
实验类型：☐ 验证 ☒ 综合 ☐ 设计 ☐ 创新 实验日期：____2021.10.23____ 实验成绩：____

一、实验项目名称

网络爬虫

二、实验目的

1. 掌握 web 数据采集技术
2. 掌握 Requests 库的使用
3. 掌握 web 信息提取技术 BeautifulSoup 库和 re 库等的使用

三、实验任务

参照房天下房价信息搜集案例和猫眼电影数据采集案例完成以下任务：

- 1、采集新闻网站的相关类别信息或豆瓣网站相关栏目信息；
- 2、对相关页面进行内容提取，生成文本集（txt 或 Excel 等形式）

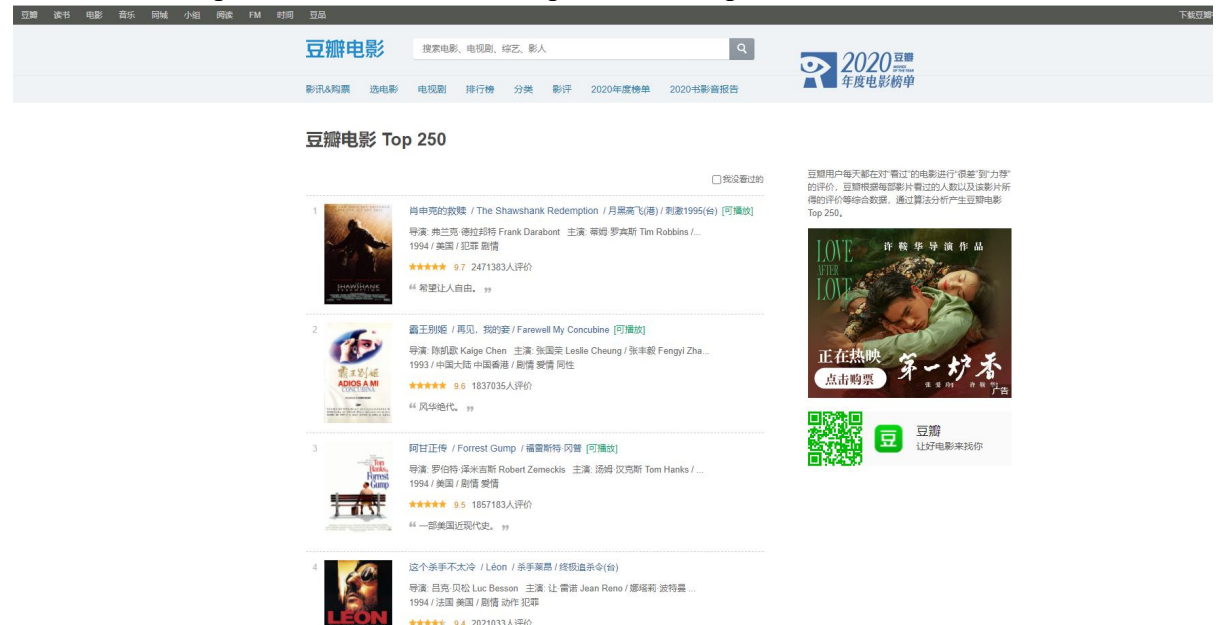
四、主要仪器设备及耗材

软件：Anaconda 或者 pycharm 等

五、实验步骤

第一步：确定采集网站和数据集合

本次实验对 <https://movie.douban.com/top250> 豆瓣 Top250 电影进行数据提取，如图。



每一页有 25 项数据，我们要找出每一项电影的电影详情链接、影片中文名、影片外国名、评分、评价数、概况、相关信息等。



以第一步电影“肖申克的救赎”为例，通过 F12 开发者模式查看网页的源代码找出相应信息的标签代码格式，便于爬虫使用正则表达式爬取相关信息。


```

16.         if hasattr(e, "reason"):
17.             print(e.reason)
18.         return response.text

```

第三步：使用正则表达式 re 和 BeautifulSoup 模块处理网页数据

搜索匹配的正则表达式，并使用 BeautifulSoup 和 findall 函数查找指定数据添加到列表 datalist 中，并返回列表数据。

```

1. #影片详情链接的规则
2. findLink = re.compile(r'<a href="(.*?)">')      #创建正则表达式对象，表示规则（字符串的模式）
3. #影片图片
4. findImgSrc = re.compile(r'<img.*src="(.*?)"',re.S)  #re.S 让换行符包含在字符串中
5. #影片片名
6. findTitle = re.compile(r'<span class="title">(.*?)</span>')
7. #影片评分
8. findRating = re.compile(r'<span class="rating_num" property="v:average">(.*?)</span>')
9. #找到评价人数
10. findJudge = re.compile(r'<span>(\d*)人评价</span>')
11. #找到概况
12. findInq = re.compile(r'<span class="inq">(.*?)</span>')
13. #找到影片的相关内容
14. findBd = re.compile(r'<p class="">(.*?)</p>',re.S)
15.
16.
17. #爬取网页函数
18. def getData(baseUrl):
19.     datalist = []
20.     for i in range(0,10):      #调用获取页面信息的函数，10次
21.         url = baseUrl + str(i*25)
22.         html = askURL(url)      #保存获取到的网页源码
23.
24.         # 2.逐一解析数据
25.         soup = BeautifulSoup(html,"html.parser")
26.         for item in soup.find_all('div',class_="item"):      #查找符合要求的字符串，形成列表
27.             #print(item)      #测试：查看电影 item 全部信息
28.             data = []      #保存一部电影的所有信息
29.             item = str(item)
30.
31.             #影片详情的链接

```

```

32.         link = re.findall(findLink,item)[0]      #re 库用来通过正则表达式查
           找指定的字符串
33.         data.append(link)                        #添加链接
34.
35.         imgSrc = re.findall(findImgSrc,item)[0]
36.         data.append(imgSrc)                      #添加图片
37.
38.         titles = re.findall(findTitle,item)      #片名可能只有一个中文名，没
           有外国名
39.         if(len(titles) == 2):
40.             ctitle = titles[0]                  #添加中文名
41.             data.append(ctitle)
42.             otitle = titles[1].replace("/","")  #去掉无关的符号
43.             data.append(otitle)                 #添加外国名
44.         else:
45.             data.append(titles[0])
46.             data.append(' ')                    #外国名字留空
47.
48.         rating = re.findall(findRating,item)[0]
49.         data.append(rating)                      #添加评分
50.
51.         judgeNum = re.findall(findJudge,item)[0]
52.         data.append(judgeNum)                    #添加评价人数
53.
54.         inq = re.findall(findInq,item)
55.         if len(inq) != 0:
56.             inq = inq[0].replace("。","")      #去掉句号
57.             data.append(inq)                    #添加概述
58.         else:
59.             data.append(" ")                    #留空
60.
61.         bd = re.findall(findBd,item)[0]
62.         bd = re.sub('<br(\s+)?/>(\s+)?',' ',bd)  #去掉<br/>
63.         bd = re.sub('/', ' ',bd)               #替换/
64.         data.append(bd.strip())                 #去掉前后的空格
65.         print(data)
66.         datalist.append(data)                  #把处理好的一部电影信息放入 datalist
67.
68.     return datalist

```

将列表中存储的数据导入到 excel 表中，使用 xlwt 模块。

```
1. # 保存数据到 excel 表中
2. def saveData(datalist,savepath):
3.     print("save...")
4.     book = xlwt.Workbook(encoding="utf-8",style_compression=0) #创建工作簿对象
5.     sheet = book.add_sheet('豆瓣电影 Top250',cell_overwrite_ok=True) #创建工作表
6.     col = ("电影详情链接","图片链接","影片中文名","影片外国名","评分","评价数","概况","相关信息")
7.     for i in range(0,8):
8.         sheet.write(0,i,col[i]) #列名
9.     for i in range(0,250):
10.        print("第%d 条" %(i+1))
11.        data = datalist[i]
12.        for j in range(0,8):
13.            sheet.write(i+1,j,data[j]) #数据
14.
15.    book.save(savepath) #保存
```

主函数调用

```
1. def main():
2.     baseurl = "https://movie.douban.com/top250?start="
3.     #1.爬取网页
4.     datalist = getData(baseurl)
5.
6.     savepath = "豆瓣电影 Top250.xls"
7.     # dbpath = "movie.db"
8.     #3.保存数据
9.     saveData(datalist,savepath)
10.
11. if __name__ == "__main__": #当程序执行时
12.     #调用函数
13.     main()
14.     askURL("https://movie.douban.com/top250?start=")
15.     #init_db("movietest.db")
16.     print("爬取完毕！")
```

六、实验数据及处理结果

(请根据实际的数据来填写)

```
spider --
[ 'https://movie.douban.com/subject/26614895/', 'https://img9.doubanio.com/view/photo/s_ratio_poster/public/p2522531945.jpg', '完美陌生人', '\xa0\xa0Perfetti sconosciuti', '1
[ 'https://movie.douban.com/subject/1768622/', 'https://img9.doubanio.com/view/photo/s_ratio_poster/public/p2441127736.jpg', '香水', '\xa0\xa0Perfume: The Story of a Murder
[ 'https://movie.douban.com/subject/1307856/', 'https://img1.doubanio.com/view/photo/s_ratio_poster/public/p2505048077.jpg', '背脊背, 脸对脸', ' ', '9.5', '82814', ' ', ' ', '导演:
[ 'https://movie.douban.com/subject/6874403/', 'https://img2.doubanio.com/view/photo/s_ratio_poster/public/p2250287733.jpg', '再次出发之纽约遇见你', '\xa0\xa0Begin Again', '8.
[ 'https://movie.douban.com/subject/5988478/', 'https://img1.doubanio.com/view/photo/s_ratio_poster/public/p1075591188.jpg', '我爱你', '\xa0\xa0그대를 사랑합니다', '9.1', '1416
[ 'https://movie.douban.com/subject/2213597/', 'https://img1.doubanio.com/view/photo/s_ratio_poster/public/p1140984198.jpg', '朗读者', '\xa0\xa0The Reader', '8.6', '407502',
[ 'https://movie.douban.com/subject/25934014/', 'https://img9.doubanio.com/view/photo/s_ratio_poster/public/p2395517414.jpg', '爱乐之城', '\xa0\xa0La La Land', '8.4', '71728
[ 'https://movie.douban.com/subject/27119724/', 'https://img9.doubanio.com/view/photo/s_ratio_poster/public/p2567198874.jpg', '小丑', '\xa0\xa0Joker', '8.7', '851883', ' ', ' ',
[ 'https://movie.douban.com/subject/25954475/', 'https://img1.doubanio.com/view/photo/s_ratio_poster/public/p2263822658.jpg', '聚焦', '\xa0\xa0Spotlight', '8.8', '241447',
[ 'https://movie.douban.com/subject/25921812/', 'https://img2.doubanio.com/view/photo/s_ratio_poster/public/p2393844761.jpg', '驴得水', ' ', '8.3', '795615', '过去的如果能让它注
[ 'https://movie.douban.com/subject/2297265/', 'https://img2.doubanio.com/view/photo/s_ratio_poster/public/p1344888983.jpg', '浪潮', '\xa0\xa0Die Welle', '8.7', '236461', ' ',
[ 'https://movie.douban.com/subject/26393561/', 'https://img9.doubanio.com/view/photo/s_ratio_poster/public/p2510956726.jpg', '小萝莉的猴神大叔', '\xa0\xa0Bajrangi Bhaijaan',
[ 'https://movie.douban.com/subject/25864085/', 'https://img2.doubanio.com/view/photo/s_ratio_poster/public/p2280097442.jpg', '火星救援', '\xa0\xa0The Martian', '8.5', '5997
[ 'https://movie.douban.com/subject/1292528/', 'https://img1.doubanio.com/view/photo/s_ratio_poster/public/p513567548.jpg', '猜火车', '\xa0\xa0Trainspotting', '8.5', '37276
[ 'https://movie.douban.com/subject/1307394/', 'https://img9.doubanio.com/view/photo/s_ratio_poster/public/p2215102596.jpg', '千年女优', '\xa0\xa0千年女優', '8.8', '201705', '
save....
第1条
第2条
第3条
第4条
第5条
第6条
第7条
第8条
第9条
第10条
第11条
第12条
第13条
第14条
第15条
第16条
第17条
第18条
第19条
第20条
第21条
第22条
第23条
第24条
第25条
第26条
第27条
第28条
第29条
第30条
第31条
第32条
第33条
第34条
第35条
第36条
第37条
第38条
第39条
第40条
第41条
第42条
第43条
第44条
第45条
第46条
第47条
第48条
第49条
第50条
第51条
第52条
第53条
第54条
第55条
第56条
第57条
第58条
第59条
第60条
第61条
第62条
第63条
第64条
第65条
第66条
第67条
第68条
第69条
第70条
第71条
第72条
第73条
第74条
第75条
第76条
第77条
第78条
第79条
第80条
第81条
第82条
第83条
第84条
第85条
第86条
第87条
第88条
第89条
第90条
第91条
第92条
第93条
第94条
第95条
第96条
第97条
第98条
第99条
第100条
第101条
第102条
第103条
第104条
第105条
第106条
第107条
第108条
第109条
第110条
第111条
第112条
第113条
第114条
第115条
第116条
第117条
第118条
第119条
第120条
第121条
第122条
第123条
第124条
第125条
第126条
第127条
第128条
第129条
第130条
第131条
第132条
第133条
第134条
第135条
第136条
第137条
第138条
第139条
第140条
第141条
第142条
第143条
第144条
第145条
第146条
第147条
第148条
第149条
第150条
第151条
第152条
第153条
第154条
第155条
第156条
第157条
第158条
第159条
第160条
第161条
第162条
第163条
第164条
第165条
第166条
第167条
第168条
第169条
第170条
第171条
第172条
第173条
第174条
第175条
第176条
第177条
第178条
第179条
第180条
第181条
第182条
第183条
第184条
第185条
第186条
第187条
第188条
第189条
第190条
第191条
第192条
第193条
第194条
第195条
第196条
第197条
第198条
第199条
第200条
第201条
第202条
第203条
第204条
第205条
第206条
第207条
第208条
第209条
第210条
第211条
第212条
第213条
第214条
第215条
第216条
第217条
第218条
第219条
第220条
第221条
第222条
第223条
第224条
第225条
第226条
第227条
第228条
第229条
第230条
第231条
第232条
第233条
第234条
第235条
第236条
第237条
第238条
第239条
第240条
第241条
第242条
第243条
第244条
第245条
第246条
第247条
第248条
第249条
第250条
爬取完毕!

Process finished with exit code 0
```

爬取完毕

Excel 表格数据

豆瓣电影Top250.xls [兼容模式] - Excel																		
文件 开始 插入 页面布局 公式 数据 审阅 视图 帮助 百度网盘 操作说明/搜索																		
剪贴板 字体 段落 对齐方式 数字 样式 单元格 编辑																		
G7 删除 格式刷 格式 自动求和 排序和筛选 查找和选择 保存 删除 格式刷 格式 自动求和 排序和筛选 查找和选择 保存																		
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
208	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	大佛普拉斯	大佛普拉斯	7	560318	人们可以登上月球，却永远无法探索人们内心的	导演	黄信尧 Hsin-yao Huang	主演	庄益增 Yung-Zheng Zhuang 竹竹汉 Chu-sh	2017	中国台湾	剧情	喜剧				
209	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	罗生门	羅生門	8	264147	人生的种种可能性	导演	黑泽明 Akira Kurosawa	主演	三船敏郎 Toshirô Mifune 京町子	1950	日本	剧情	犯罪	悬疑			
210	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	终结者2：审判日	Terminator 2: Judgment Day	8	297176	少见的超越前作的续集，动作片中的经典	导演	詹姆斯·卡梅隆 James Cameron	主演	阿诺·施瓦辛格 Arnold Schwar	1991	美国	动作	科幻				
211	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	杨善庆少年杀人事件	杨善庆少年殺人事件	8	229229	杨善庆少年杀人事件	导演	杨德昌 Edward Yang	主演	张震 Chen Chang 杨静怡 Lisa Yang	1991	中国台湾	剧情	犯罪				
212	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	房间	Room	8	229998	被幽禁的岁月，被囚禁的生命，被禁锢的灵魂	导演	伦尼·阿布拉汉森 Lenny Abrahamson	主演	希瑟·拉姆森 Brie Larson	2015	爱尔兰 加拿大 英国 美国	剧情	家庭				
213	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	青蛇	青蛇	8	145465	人生如此，浮生如斯，佛人言，花彼岸，此生情长	导演	徐克 Hark Tsui	主演	张曼玉 Maggie Cheung 王祖贤 Joey Wang	1993	中国香港 中国大陆	剧情	爱情	奇幻			
214	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	源代码	Source Code	8	705504	邓肯·琼斯	导演	邓肯·琼斯 Duncan Jones	主演	杰克·吉伦哈尔 Jake Gyllenhaal	2011	美国 加拿大	科幻	悬疑	惊悚			
215	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	无间道	無間道	8	372204	无间道	导演	刘伟强 麦兆辉	主演	刘德华 梁朝伟 黄秋生 曾志伟	2003	中国香港	剧情	犯罪	惊悚			
216	https://movie.douban.com/subject/https://img1.doubanio.com/view/photo/large/public/p250.jpg	新龙门客栈	新龍門客棧	8	579111	龙门客栈	导演	李惠民 Raymond Lee	主演	张曼玉 Maggie Cheung 林青霞 Brigitte	1992	中国香港 中国大陆	动作	爱情	武侠			
217	https://movie.douban.com/subject/https://img1.doubanio.com/view/photo/large/public/p250.jpg	东京教父	東京ゴッドファーザーズ	9	163711	东京教父	导演	今敏 Satoshi Kon	主演	江守彻 Toru Emori 梅田道弘 Yoshiaki Ume	2003	日本	剧情	动画	冒险			
218	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	狂暴之路	Mad Max: Fury Road	8	448880	狂暴之路	导演	乔治·米勒 George Miller	主演	汤姆·哈迪 Tom Hardy 查理兹·塞隆	2015	澳大利亚 美国	动作	科幻	冒险			
219	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	波西米亚狂想曲	Bohemian Rhapsody	8	499393	波西米亚狂想曲	导演	布莱恩·辛格 Bryan Singer	主演	拉米·马雷克 Rami Malek	2018	英国 美国	剧情	传记	音乐			
220	https://movie.douban.com/subject/https://img1.doubanio.com/view/photo/large/public/p250.jpg	温斯顿	Waterloo Bridge	8	257786	温斯顿	导演	瓦茨·李维 Wally Pfister	主演	凯拉·奈特莉 Keira Knightley 詹姆斯·麦卡沃伊	2010	英国	剧情	战争				
221	https://movie.douban.com/subject/https://img1.doubanio.com/view/photo/large/public/p250.jpg	无耻混蛋	Inglourious Basterds	8	407259	无耻混蛋	导演	昆汀·塔伦蒂诺 Quentin Tarantino	主演	布拉德·皮特 Brad Pitt 梅丽尔·斯特里普	2009	德国 美国	剧情	犯罪				
222	https://movie.douban.com/subject/https://img1.doubanio.com/view/photo/large/public/p250.jpg	步履不停	歩いても 歩いても	8	233331	步履不停	导演	是枝裕和 Hirokazu Koreeda	主演	阿部宽 Hiroshi Abe 夏川结衣	2008	日本	剧情	家庭				
223	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	血钻	Blood Diamond	8	316800	血钻	导演	爱德华·兹威克 Edward Zwick	主演	莱昂纳多·迪卡普里奥 Leonardo	2006	美国 德国 英国	剧情	惊悚	冒险			
224	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	盗梦空间	Inception	8	458963	盗梦空间	导演	克里斯托弗·诺兰 Christopher Nolan	主演	莱昂纳多·迪卡普里奥 Leonardo	2010	美国 英国	动作	科幻	惊悚			
225	https://movie.douban.com/subject/https://img1.doubanio.com/view/photo/large/public/p250.jpg	战争之王	Lord of War	8	250222	战争之王	导演	安德鲁·尼克 Andrew Niccol	主演	尼古拉斯·凯奇 Nicolas Cage	2005	法国 德国 美国	剧情	犯罪				
226	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	千钧一发	Gattaca	8	233645	千钧一发	导演	安德鲁·尼克 Andrew Niccol	主演	伊桑·霍克 Ethan Hawke	1997	美国	剧情	科幻	惊悚			
227	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	盗梦空间2	The Bourne Supremacy	8	305395	盗梦空间2	导演	保罗·格林格拉斯 Paul Greengrass	主演	马特·达蒙 Matt Damon	2004	美国 德国	动作	惊悚	惊悚			
228	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	盗梦空间	The Bourne Supremacy	8	305395	盗梦空间	导演	保罗·格林格拉斯 Paul Greengrass	主演	马特·达蒙 Matt Damon	2004	美国 德国	动作	惊悚	惊悚			
229	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	盗梦空间	The Bourne Supremacy	8	305395	盗梦空间	导演	保罗·格林格拉斯 Paul Greengrass	主演	马特·达蒙 Matt Damon	2004	美国 德国	动作	惊悚	惊悚			
230	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	海洋	Océans	9	143408	海洋	导演	雅克·贝汉 Jacques Perrin 雅克·克卢佐 Jacques Clouzot	主演	雅克·贝汉 Jacques Perrin 雅克·克卢佐 Jacques Clouzot	2009	法国 瑞士 西班牙 美国 阿联酋 摩纳哥	纪录片	自然				
231	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	黑客帝国2：重装上阵	The Matrix Reloaded	8	333668	黑客帝国2：重装上阵	导演	拉娜·沃卓斯基 Lana Wachowski 莉莉·沃卓斯基	主演	基努·里维斯 Keanu Reeves 凯瑞-安·莫斯	2003	美国	动作	科幻				
232	https://movie.douban.com/subject/https://img1.doubanio.com/view/photo/large/public/p250.jpg	星上的痕迹	星の上的痕迹	8	401535	星上的痕迹	导演	高田浩吉 Hayao Miyazaki	主演	宫崎骏 Hayao Miyazaki	2008	日本	动画	奇幻	冒险			
233	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	阿飞正传	阿飛正傳	8	330424	阿飞正传	导演	王家卫 Kar Wai Wong	主演	张国荣 Leslie Cheung 张曼玉 Maggie C	1990	中国香港 日本	剧情	爱情				
234	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	穿越时空的少女	時をかける少女	8	336454	穿越时空的少女	导演	细田守 Mamoru Hosoda	主演	仲里依纱 Risa Nakai 石田卓也	2006	日本	剧情	爱情	科幻	动画		
235	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	谍影重重	The Bourne Identity	8	365850	谍影重重	导演	道格·里曼 Doug Liman	主演	马特·达蒙 Matt Damon 弗兰卡·波坦	2002	美国 德国 法国	动作	惊悚	惊悚			
236	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	地球上的星星	Taare Zameen Par	8	179335	地球上的星星	导演	阿米尔·汗 Amir Khan	主演	达席尔·萨法瑞 Darshab Saty	2007	印度	剧情	儿童	家庭			
237	https://movie.douban.com/subject/https://img1.doubanio.com/view/photo/large/public/p250.jpg	完美陌生人	Perfetti sconosciuti	8	465392	完美陌生人	导演	保罗·格诺维塞 Paolo Genovese	主演	马可·贾利尼 Marco Giallin	2016	意大利	剧情	喜剧				
238	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	香水	Perfume: The Story of a Murderer	8	400697	香水	导演	汤姆·提克威 Tom Tykwer	主演	本·卫肖 Ben Whishaw 艾文·布莱纳	2006	德国 法国 西班牙 英国 比利时	剧情	犯罪				
239	https://movie.douban.com/subject/https://img1.doubanio.com/view/photo/large/public/p250.jpg	曾孝舜，你对她	Begin Again	8	28214	曾孝舜，你对她	导演	贾斯汀·汉密尔顿 Justin Hamilton	主演	贾斯汀·汉密尔顿 Justin Hamilton	2013	美国	剧情	音乐				
240	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	再次出发之伯约克	Begin Again	8	28214	再次出发之伯约克	导演	贾斯汀·汉密尔顿 Justin Hamilton	主演	贾斯汀·汉密尔顿 Justin Hamilton	2013	美国	剧情	音乐				
241	https://movie.douban.com/subject/https://img1.doubanio.com/view/photo/large/public/p250.jpg	再次出发之伯约克	Begin Again	8	28214	再次出发之伯约克	导演	贾斯汀·汉密尔顿 Justin Hamilton	主演	贾斯汀·汉密尔顿 Justin Hamilton	2013	美国	剧情	音乐				
242	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	再次出发之伯约克	Begin Again	8	28214	再次出发之伯约克	导演	贾斯汀·汉密尔顿 Justin Hamilton	主演	贾斯汀·汉密尔顿 Justin Hamilton	2013	美国	剧情	音乐				
243	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	再次出发之伯约克	Begin Again	8	28214	再次出发之伯约克	导演	贾斯汀·汉密尔顿 Justin Hamilton	主演	贾斯汀·汉密尔顿 Justin Hamilton	2013	美国	剧情	音乐				
244	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	再次出发之伯约克	Begin Again	8	28214	再次出发之伯约克	导演	贾斯汀·汉密尔顿 Justin Hamilton	主演	贾斯汀·汉密尔顿 Justin Hamilton	2013	美国	剧情	音乐				
245	https://movie.douban.com/subject/https://img1.doubanio.com/view/photo/large/public/p250.jpg	再次出发之伯约克	Begin Again	8	28214	再次出发之伯约克	导演	贾斯汀·汉密尔顿 Justin Hamilton	主演	贾斯汀·汉密尔顿 Justin Hamilton	2013	美国	剧情	音乐				
246	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	再次出发之伯约克	Begin Again	8	28214	再次出发之伯约克	导演	贾斯汀·汉密尔顿 Justin Hamilton	主演	贾斯汀·汉密尔顿 Justin Hamilton	2013	美国	剧情	音乐				
247	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	再次出发之伯约克	Begin Again	8	28214	再次出发之伯约克	导演	贾斯汀·汉密尔顿 Justin Hamilton	主演	贾斯汀·汉密尔顿 Justin Hamilton	2013	美国	剧情	音乐				
248	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	少罗的罗神大坂	Bajrang Bhajjan	8	428275	少罗的罗神大坂	导演	卡比汗 Kabir Khan	主演	伊尔汗 Salman Khan 帕莎·帕西	2015	印度	剧情	喜剧	动作			
249	https://movie.douban.com/subject/https://img2.doubanio.com/view/photo/large/public/p250.jpg	罗神大坂	The Marican	8	599740	罗神大坂	导演	卡比汗 Kabir Khan	主演	伊尔汗 Salman Khan 帕莎·帕西	2015	印度	剧情	喜剧	动作			
250	https://movie.douban.com/subject/https://img1.doubanio.com/view/photo/large/public/p250.jpg	罗神大坂	The Marican	8	599740	罗神大坂	导演	卡比汗 Kabir Khan	主演	伊尔汗 Salman Khan 帕莎·帕西	2015	印度	剧情	喜剧	动作			
251	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	罗神大坂	The Marican	8	599740	罗神大坂	导演	卡比汗 Kabir Khan	主演	伊尔汗 Salman Khan 帕莎·帕西	2015	印度	剧情	喜剧	动作			
252	https://movie.douban.com/subject/https://img3.doubanio.com/view/photo/large/public/p250.jpg	罗神大坂	The Marican	8	599740	罗神大坂	导演	卡比汗 Kabir Khan	主演	伊尔汗 Salman Khan 帕莎·帕西	2015	印度	剧情	喜剧	动作			

七、思考讨论题或体会或对改进实验的建议

(请根据实际的数据来填写)

八、参考资料

[1] [美] 阿曼多·凡丹戈 (Armando Fandango) 著, 韩波 译. Python 数据分析(第2版). 人民邮电出版社.2018.6

[2] 嵩天, 礼欣, 黄天羽 著.Python 语言程序设计基础 (第2版). 高等教育出版社.2017.2